

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-642-734C-4

Query Match 3.7%; Score 115; DB 2; Length 3567;
Best Local Similarity 22.8%; Pred. No. 0.71; Indels 174; Gaps 27;
Matches 133; Conservative 50; Mismatches 226;

```

21 AFAVTEVTDELLANPPAGEW-----ISYQONQENTRHSPLTQITTEVNGQLQIYW 71
2911 AFAVIAVTDRLRLARVDDG-WSDADAAPVIAV-----TTAHYALHDLAGLAGOSVLIH 2964
72 ARGMPKQVQVYPLIHGCVYIANPGDVIOAIDAKTGDLIWEHRQLPNIATLNSGEPT 131
2965 AAAGVGMAAVALARRAGAEVLATAG-----PAKHG-----TLRALG----- 3001
132 RGMALYGTNVYFVSMNHLVALDTATGQVTEVDYR--GQGEDMVNS-----SGPIV 181
3002 -----LDEHIASSRETEGFARKFRERTGKRGVDVYVLSLIGELLDESADLL 3047
182 A-NGVIV-AGST-----CQSPGCFVSGHDSATGELMRNFIPRAGEGETWG 230
3048 AEDGVFVEMKTDLRDAGDFRGARYAPFDLGEAG--DRLGLILEVYVGLLAGELD----- 3101
231 NDYERAMWTGAM---GQITVDPTNLVHYGSTAV--GPASETORGT---PGTLYGTNTR 282
3102 -----RLPVSAMELGSAAPALQHMRSRGRHVAKLVLPAPVDPDGYLLITGGR--GTIGR 3154
283 FAVRPDTGEIYWRHQT--PRDNMDQECTFEMMTVNVDPQSTMEGLQSTINPNAAT----- 337
3155 LLARLIVTEHGVRLHLIVSRGADAPGSDDELRAETEDLGASAEIACDTPADRLALSLLD 3214
338 GERAVLTGVPCKTGMQFPAETGEFLMARPTNYONNIESIDENGIYT--VNEDAILKELD 336
3215 GLPRLTGVVHAAGVL-----ADGLVTSIDEPRAVQVYLR 3248
397 VEYDVCPTFLGGRDWPSSALNPDGSIYFIPLNNVCYDMAVDOEFTSMDYNTSNV----- 452
3249 AKVDAA-----WNLHELTAANTGLSFVLFSSAASVLA-----GPGQGYAAANESLNA 3296
453 -----TKLPKG-----DMGRIDAIDISTGRITLMSVERAANYSP 488
3297 LLAALRTKGLPAKALGWMQASMTSGLDRIAR-----TGVAALPTERALLFDS 3349
489 VLTGGGVLFNGGTDYFRALSOETGETL--WQTRLATVASGOA 530
3350 ALRRGGEVVFPLINSALRRAEFVPEVLRGMVRAKLLRAAQA 3392

```

RESULT 7

US-08-439-009A-4
Sequence 4, Application US/08439009A
Patent No. 6004787

GENERAL INFORMATION:

APPLICANT: Donadio, S
APPLICANT: Katz, L
APPLICANT: McAlpine, J B
TITLE OF INVENTION: Method of Directing Biosynthesis of
SPECIFIC POLYPEPTIDES
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Steven F. Weinstein
STREET: Abbott Laboratories D377/APD-2 One Abbott
CITY: Abbott Park
STATE: IL
COUNTRY: US
ZIP: 60064-3500

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/439,009A
FILING DATE: 11-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Casuto, Dianne
REGISTRATION NUMBER: 40,943
REFERENCE/DOCKET NUMBER: 4952, US.D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847-938-3137
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3567 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-439-009A-4

Query Match 3.7%; Score 115; DB 3; Length 3567;
Best Local Similarity 22.8%; Pred. No. 0.71;
Matches 133; Conservative 50; Mismatches 226; Indels 174; Gaps 27;

```

21 AFAVTEVTDELLANPPAGEW-----ISYQONQENTRHSPLTQITTEVNGQLQIYW 71
2911 AFAVIAVTDRLRLARVDDG-WSDADAAPVIAV-----TTAHYALHDLAGLAGOSVLIH 2964
72 ARGMPKQVQVYPLIHGCVYIANPGDVIOAIDAKTGDLIWEHRQLPNIATLNSGEPT 131
2965 AAAGVGMAAVALARRAGAEVLATAG-----PAKHG-----TLRALG----- 3001
132 RGMALYGTNVYFVSMNHLVALDTATGQVTEVDYR--GQGEDMVNS-----SGPIV 181
3002 -----LDEHIASSRETEGFARKFRERTGKRGVDVYVLSLIGELLDESADLL 3047
182 A-NGVIV-AGST-----CQSPGCFVSGHDSATGELMRNFIPRAGEGETWG 230
3048 AEDGVFVEMKTDLRDAGDFRGARYAPFDLGEAG--DRLGLILEVYVGLLAGELD----- 3101
231 NDYERAMWTGAM---GQITVDPTNLVHYGSTAV--GPASETORGT---PGTLYGTNTR 282
3102 -----RLPVSAMELGSAAPALQHMRSRGRHVAKLVLPAPVDPDGYLLITGGR--GTIGR 3154
283 FAVRPDTGEIYWRHQT--PRDNMDQECTFEMMTVNVDPQSTMEGLQSTINPNAAT----- 337
3155 LLARLIVTEHGVRLHLIVSRGADAPGSDDELRAETEDLGASAEIACDTPADRLALSLLD 3214
338 GERAVLTGVPCKTGMQFPAETGEFLMARPTNYONNIESIDENGIYT--VNEDAILKELD 336
3215 GLPRLTGVVHAAGVL-----ADGLVTSIDEPRAVQVYLR 3248
397 VEYDVCPTFLGGRDWPSSALNPDGSIYFIPLNNVCYDMAVDOEFTSMDYNTSNV----- 452
3249 AKVDAA-----WNLHELTAANTGLSFVLFSSAASVLA-----GPGQGYAAANESLNA 3296
453 -----TKLPKG-----DMGRIDAIDISTGRITLMSVERAANYSP 488
3297 LLAALRTKGLPAKALGWMQASMTSGLDRIAR-----TGVAALPTERALLFDS 3349
489 VLTGGGVLFNGGTDYFRALSOETGETL--WQTRLATVASGOA 530
3350 ALRRGGEVVFPLINSALRRAEFVPEVLRGMVRAKLLRAAQA 3392

```

RESULT 8

US-08-700-651-5
Sequence 5, Application US/08700651B
Patent No. 6015882

GENERAL INFORMATION:

APPLICANT: PETERSEN, CAROLYN
APPLICANT: LEECH, JAMES

APPLICANT: NELSON, RICHARD, C.
 APPLICANT: GUT, JIRI
 TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS
 TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
 TITLE OF INVENTION: INFECTIONS
 FILE REFERENCE: 480.19-4(HV)
 CURRENT APPLICATION NUMBER: US/08/700,651B
 EARLIER FILING DATE: 1997-08-14
 EARLIER APPLICATION NUMBER: 08/415,751
 EARLIER FILING DATE: 1995-04-03
 NUMBER OF SEQ ID NOS: 15
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 5
 LENGTH: 1721
 TYPE: PRT
 ORGANISM: Cryptosporidium parvum
 US-08-700-651-5

Query Match 3.7%; Score 114.5; DB 3; Length 1721;
 Best Local Similarity 20.2%; Pred. No. 0.24;
 Matches 140; Conservative 82; Mismatches 227; Indels 245; Gaps 38;

25 VTPVDELAMPAGEMWISYQONE-----NTRHSPLOITTEVNGQLQVWARGMOP 77
 979 IDPTTG-LPEPPTGHLINPTNNMTDSSPAGAYKAYVNSGIKTDV-----YGLPV 1029
 78 GKVOVTPLIHGVVWVLANPDPVIOAIDAKTGDDIWEHRQDLPNATINSCEPTRGALY 137
 1030 GEIGLGP-----KDPGSDI-PFNPSTGELY-----DSTGKPIN 1062
 138 GTNVEYSDMNLVALDPTAGVQVTEVDVROGEDMNSSGPIYANGVIVAGSTCQSPF 197
 1063 NSTAGIVSGKRGKLPIDENGNL-ED-----PSTMLPIDGNNOQLNPET----- 1105
 198 GCEFSGSDSAT-----GEELMRNFIP--RAGEGDETWINDYEMARMGAMQIITYDPV 250
 1106 NSTVSGSTGTTKPKKPIPVNGGVVDEBAKQADK--GKD--GLIIVPTNSINKDPV 1160
 251 TNLVHGSTA--VGPASETORGTG-----GTLCGTNRRAVRP 287
 1161 TMOQYMTGNTIINP--ETCKVPIGSLPGLANPSPNTPOQDITETGKPYDVTGLPYDP 1218
 288 DTGEIYVRRHQTLPDND--WDQCTEEMVYN-----VDVQ--PSTEMEGLQS----- 330
 1219 STGEIIDPATKLPKIPGSVAGDEILTEVLNITDEVVGLPIDETGLPRDPVSGLPQLPNG 1278
 331 --INP-----NAATG-----ERRVLTGVPCKTGT--MMQPD-----AETG 361
 1279 TLVDPSSKRRIPGSHSGFTIGTSGEOSHEDPSTGKPLDPNTGLHPFDEDSGLINPETG 1338
 362 EFLMARDTNY-----ONMIESIDENGIVTYNE-----PAI----- 391
 1339 DKLOGSHSGTFMPVPGKPGGENGIMTPEQIIEAL--MLPITSNEVINISPRSSSDVPR 1396
 392 -----LKEIDVEVDCPTFLG-----RDWPSAAL--NPDSGIYFIPLN 428
 1397 PNTTMMNKISGQTYQVDAKKTIIASAASVIAHTALGPTDPTTGLPSPSGLPIFIEG 1456
 429 NNCYDMAVDO-----EFTSMVYNTSNVTKLPKPKDMIGRIDAIDISTGRILMSVERAA 483
 1457 NVLVDPQTEQJIKGVPIVSLVYKKNVTEAAYGL-----PVDPKTGPIIDPI----- 1505
 484 ANSPVLSTG-----GGVLFNGGTDTRYFRA-----LSOETGETLMQT----- 520
 1506 -SYLFPKANGCELIDPISGKYFGSGLAGFISCAKASQSSKSDSGNPIDPSTMPDPKGG 1564
 521 RATAVASGAISEYVDGMQYVAIAG-----GGV 548
 1565 KLIDPESGAIIDNSVSGV-FATVPGTAAPKKGCV 1597

RESULT 9

US-08-219-262B-10
 Sequence 10, Application US/08219262B
 Patent No. 5788970
 GENERAL INFORMATION:
 APPLICANT: VAKHARIA, VIKRAM
 APPLICANT: SNYDER, DAVID B
 APPLICANT: MENDEL-WHERSAT, STEPHANIE A
 TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS
 TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
 STREET: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR
 CITY: ARLINGTON
 STATE: VIRGINIA
 COUNTRY: USA
 ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/219,262B
 FILING DATE: 29-MAR-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: OBLON, NORMAN F
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 2747-047-27
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 413-2220
 TELEFAX: (703) 413-3000
 TELEX: 248855 OPAT UR
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1012 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:
 ORGANISM: Infectious bursal disease virus
 STRAIN: OH
 US-08-219-262B-10

Query Match 3.7%; Score 113.5; DB 1; Length 1012;
 Best Local Similarity 23.0%; Pred. No. 0.12;
 Matches 94; Conservative 49; Mismatches 139; Indels 127; Gaps 22;

250 VTNLVHGSTAV-----GPAS-----ETORGTGCTLYGTNTRF 283
 1 MTNLMHTQOIVPFIISLMLPTGPASIPDTELEKHTLRSTSYNLTVDGSGGLYVF 60
 284 AVRPDTGIVVRRHQTLPDND--WDQCTEEMVYN-----VDVQ--PSTEMEGLQS----- 330
 61 PGFP--GSVAHAHTLOSNGSYOPDQ-----MLTAAONLPSVYNYCHLVRSRLTVRSSTL 113
 325 MEGLOSTNP--NAATGRRVLTGVPCKTGTMMQPDATGTEFLMARDTNYNMIEST----- 378
 114 PGVYALNGTINAVT-----FOGSLSEL--TDSYNGLSASATANIN 152
 379 DENGIIVVNEDAILKELDEVDCPTFLAGRDWPSAALNPDSGIYFIPLNNVC----- 431
 153 DKIGNVLVGGGVFLSPTSYLSYRLGD--PIPAAGLDP-----KLMAOTDSSDRR 204
 432 -YDMAMVDO--EF-----TSMVYNTSNVTKLPKPKDMIGRIDAIDISTGRILM- 477
 205 VYTVTADEVQSSQLIPSGVKTTLFTANIDALTSVGGELIFSGVTHSIEVDVTYF 264
 478 -----SVERAANSPVLSTGGVLFN--GGTDTRYFRAISOETGETLMQTRLATVAS 527

Db 265 IGPDETEYVAKAVADPGLTTGTNNLVPFNLGGPTSETPTPTSMKLEVTYKRGCT--A 322
QY 528 GOAISYEVDMQYVVALAGGVSYGSLN--SALAGERVDTAIGNAVY 574
Db 323 GDPISWTVSGTLAVTVGG--NYPGALRPVTLVAYERV--AAGSVTV 366

RESULT 10

US-09-031-655-10
Sequence 10, Application US/09031655
Patent No. 6017759
GENERAL INFORMATION:
APPLICANT: VAKHARIA, VIKRAM
APPLICANT: SNYDER, DAVID B
APPLICANT: MENDEL-WHERSAT, STEPHANIE A
TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS
TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED
TITLE OF INVENTION: THERON
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
STREET: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031,655
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/219,262
FILING DATE: 29-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 2747-047-27
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1012 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Infectious bursal disease virus
STRAIN: OH
US-09-031-655-10

Query Match 3.7%; Score 113.5; DB 3; Length 1012;
Best Local Similarity 23.0%; Pred. No. 0.12;
Matches 94; Conservative 49; Mismatches 139; Indels 127; Gaps 22;

QY 250 VNNLVHYSSTAY-----GPAS-----EFGRTPGGTLYGTNTRF 283
Db 1 WNNLMDHQOIVPTKSLMPTGPASIPDDTLEKHTLRSEISTYNLTVGDTGSLIYVF 60
QY 284 AVRPDTGEIVMRHQTLPD--NMDQECTFEMAVT-----NVDVQPSST 324
Db 61 PGP--GSVVAHAHTLQSGSYQFDQ---MLTFAQNLVPSVNYCRVLSRLTVRSSTL 113
QY 325 MEGLOSINP--NAATGERAVLTGVECKTGTMTQDAEIGEFIMADTYQNNIESI----- 378

Db 114 PGYVALNNTINAV-----FGSLSEL---TDYSYNGLSATANIN 152
QY 379 DENGIVTVNEDAILKEDVEYDVCPTFLGGRWPSAALNPDSGITFIFLNNVC----- 431
Db 153 DKIGNVLVEGVTVLSLPTSYDLSTYVRIGD--PIPAAGIDP-----KIMATCDSSDRP 204
QY 432 -YDMAAVDO-EF-----TSMDEVNTSNVTKLPKPKDMI-GRIDALDISTGRTLM- 477
Db 205 VYTVTAADDEYQFSSQLIPSGVKTTLFTANIDALHLSVGGELIFSQVTHLSIEVDVTVIF 264
QY 478 -----SVERAANYSPVLSTGGVLFN--GGTDYFRALSOETGETLMQFRLATVAS 527
Db 265 IGPDETEYVAKAVADPGLTTGTNNLVPFNLGGPTSETPTPTSMKLEVTYKRGCT--A 322
QY 528 GOAISYEVDMQYVVALAGGVSYGSLN--SALAGERVDTAIGNAVY 574
Db 323 GDPISWTVSGTLAVTVGG--NYPGALRPVTLVAYERV--AAGSVTV 366

RESULT 11

US-08-542-003-6
Sequence 6, Application US/08542003
Patent No. 5864013
GENERAL INFORMATION:
APPLICANT: Goldberg, Edward B.
TITLE OF INVENTION: MATERIALS FOR THE PRODUCTION OF
TITLE OF INVENTION: NANOMETER STRUCTURES AND USE THEREOF
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie and Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: US
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/542,003
FILING DATE: 13-OCT-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 8471-0005-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 790-8664
TELEX: 66441 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1026 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Bacteriophage T4
IMMEDIATE SOURCE:
CLONE: p37 amino acid
US-08-542-003-6

Query Match 3.6%; Score 112.5; DB 2; Length 1026;
Best Local Similarity 20.3%; Pred. No. 0.16;
Matches 145; Conservative 74; Mismatches 281; Indels 215; Gaps 36;

QY 8 WASAGA-LALLAFAAQVTVVTEDELLANPPAGEWISYGOENRYRHSPLTQTTEWVG 66
Db 351 WIMPGTNALLSVQTQAD-----NNNAGD-----GOTHIGYNAGKMHYFRCTGQ 396


```

US-08-928-361B-30
; Sequence 30, Application US/08928361B
; Patent No. 6071518
; GENERAL INFORMATION:
; APPLICANT: Petersen, Carolyn
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: SPECIES INFECTIONS
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETERS, VERNY, JONES & BIRKA
; STREET: 385 Sherman Avenue, Suite 6
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-1840
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,361B
; FILING DATE: 12-SEP-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,062
; FILING DATE: 13-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: VERNY, HANA
; REGISTRATION NUMBER: 30,518
; REFERENCE/DOCKET NUMBER: 480,76-1(HV)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-1677
; TELEFAX: 650-324-1678
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1043 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-928-361B-30

```

Query Match 3.58; Score 108.5; DB 3; Length 1043;
 Best Local Similarity 20.3%; Pred. No. 0.38;
 Matches 141; Conservative 82; Mismatches 226; Indels 245; Gaps 39;

```

QY 25 VTPYDELALNPPAGEMISYQONE-----NYRHSPLIQITTEVNGQLVLRGMQRP 77
DB 301 IDPTTG-LPENPPTGHLINPTNNMTDSSFAGAKYAVNSGKIDNV-----YGLPV 351
QY 78 GKVQVTPLIHGVNVLNAPGDVIOAIDAKTGLDILWEHRQLPNLTIATLNSGEPTRGALY 137
DB 352 GEITGLP-----KDRGSDI-PFNSTIGELY-----DPTGKPIN 384
QY 138 GTNAVYVSMNDHLVALDPTANGVTFVDYDRGGEDMVSNSGPIYANGVYAGSTQYSPF 197
DB 385 NSTAGIVSGKGLPIEDENGNTL-FD-----PSTNIPIGNNQLVNPET----- 427
QY 198 GCFVSGHDSAT-----GELMRNFIP--RAGEGDETFWGDYDARMMTGAMQGITYPV 250
DB 428 NSTVSGSTGTTKRPKGPVGVGGVVDDEAKDADK--GKD--GLIVPTNSINKDPV 482
QY 251 TNLVHYGSTA--VGPASETORGTPG-----GLYGTNTRFAVRP 287
DB 483 TNLQVSNNTGNINP--ETGKVIQSGSLPGSLNYSFNPQOTDITLTKPVDVTVGLPYDP 540
QY 288 DTGETVWRHQTLPNDN--WDQECTFEMMYTN-----VDVQ--PSTMEGLOS----- 330
DB 541 STGEIIDPATKLPJPGSVAGDEILTEVLNITTDVETGLPIDLETGLPRDPVSGILPQLPNG 600

```

```

QY 331 --INP-----NAATG-----ERRVLTGVPCKTGT--MMQFD-----AETG 361
DB 601 TLVDPSSNKKPIPGSGHSLNGTSGEQSHKDPSTKPLDPNTGILRPDPBDSGLINPRTG 660
QY 362 EFLMARDTNY-----QNMIESIDENGIVTVN-----DAI--- 391
DB 661 DLQSGHSTFMPVPGKPGENGIMTPQDLEAL--NKLPTSNBNVISPRSSDVPDR 718
QY 392 -----LKELDVEYD-----VCPTFLG--GRDMSAL--NPDGIFIFLN 428
DB 719 PNTWMNKISGOTYVDGKRTIGSAASVIRHATLPTQDPTGLPSPDSTGLPIPGF 778
QY 429 NYCVDMAVDO-----EFTSMVNTSNVTKLPCKMDIGRIDAIDISTGRTIMSVERRA 483
DB 779 NVLVPQRTGEQIKGSVPYSLYVKEKNIVTEAAYGL-----PYDPKGTGPIDPI----- 827
QY 484 ANSPVSLTG-----GGVLFNGGIDRYFRA-----LSQETGETLMQT----- 520
DB 828 -SYLPPAKNGELIDPISGKVFSGSIAGISGRAGSQSKSDSGNPDIPSTNMPYPKGG 886
QY 521 RLATVASGQAISEYDGMQYVLAG-----GGV 548
DB 887 KLIDPESGAIIDNSVSGV-FATVPGTAAPKKGv 919

```

RESULT 14
 US-08-928-361B-6
 ; Sequence 6, Application US/08928361B
 ; Patent No. 6071518
 ; GENERAL INFORMATION:
 ; APPLICANT: Petersen, Carolyn
 ; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
 ; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
 ; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
 ; NUMBER OF SEQUENCES: 30
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: PETERS, VERNY, JONES & BIRKA
 ; STREET: 385 Sherman Avenue, Suite 6
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94306-1840
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/928,361B
 ; FILING DATE: 12-SEP-1997
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/026,062
 ; FILING DATE: 13-SEP-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: VERNY, HANA
 ; REGISTRATION NUMBER: 30,518
 ; REFERENCE/DOCKET NUMBER: 480,76-1(HV)
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 650-324-1677
 ; TELEFAX: 650-324-1678
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1721 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-928-361B-6

MEDIUM TYPE: Floppy disk

[illegible]

Db 827 ---SYLPPAKNGELIDPISGKYFSGSIAGFISGKAGSQSKSSDESGNPIDPSTNMPYDPK 883
OY 521 --RLATVASGOAISYEVDGMQYVAIAG-----GGV 548
 : | | | | : : | | |
Db 884 TGLIDPESGIAIDNSVSGV-FATVPGTAAPKKGCV 918

Search completed: August 8, 2001, 19:41:26
Job time: 83 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 2001, 19:40:03 ; Search time 19.51 Seconds
(without alignments)
2260.639 Million cell updates/sec

Title: US-08-934-506a-5

Perfect score: 3089
Sequence: 1 MKPTSLIMASAGALALLAAP.....GERVDSTALGNVYFALPQ 579

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR:68:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	583	18.9	708	2 S52317	quinohemoprotein e
2	564.5	18.3	738	2 S14370	alcohol dehydrogen
3	561	18.2	742	2 A49340	alcohol dehydrogen
4	547.5	17.7	742	2 JS0326	alcohol dehydrogen
5	505	16.3	623	2 B83399	quinoprotein alcoh
6	468.5	15.2	626	2 JU0706	alcohol dehydrogen
7	439.5	14.2	573	2 S68591	methanol dehydroge
8	392.5	12.7	796	1 JV0107	glucose dehydrogen
9	392.5	12.7	796	2 H85495	glucose dehydrogen
10	390	12.6	801	1 S00943	glucose dehydrogen
11	340	11.0	808	1 QPKEX	glucose dehydrogen
12	327	10.6	803	2 F83360	glucose dehydrogen
13	313.5	10.1	809	2 A55547	quinolate-shikimate
14	276.5	9.0	639	2 JC4881	polyvinyl-alcohol
15	221.5	7.2	221	2 A41378	hypothetical prote
16	167.5	5.4	407	2 H69064	serine/threonine p
17	167.5	5.4	524	2 A82580	polyvinylalcohol d
18	159.5	5.0	392	2 B85895	probable dehydroge
19	155.5	5.0	392	2 G65027	hypothetical prote
20	148.5	4.8	668	2 CJ5264	probable serine/th
21	145	4.7	1645	2 JN0896	crystalline surfac
22	142	4.6	943	2 JC4081	sucrase/fructanase
23	138.5	4.5	424	2 T29137	hypothetical prote
24	132.5	4.3	13055	2 T16580	hypothetical prote
25	132	4.3	407	2 C82804	conserved hypothet
26	128	4.1	1365	2 A41463	glucosyltransferas
27	126.5	4.1	5188	2 B85547	probable RTX fami
28	126	4.1	613	2 F69424	conserved hypothet
29	126	4.1	955	2 G64866	probable membrane

30	124.5	4.0	1588	2 A86036	probable adhesin Z
31	122.5	4.0	547	2 H85699	partial probable a
32	122.5	4.0	644	1 I40712	endo-1,4-beta-xyla
33	122.5	4.0	799	2 T48889	serine/threonine p
34	122	3.9	386	2 A82284	conserved hypothet
35	121.5	3.9	603	2 F72237	conserved hypothet
36	121.5	3.9	827	2 F4512	hypothetical prote
37	121.5	3.9	1441	2 B86807	hypothetical prote
38	120.5	3.9	1197	2 D82696	conserved hypothet
39	120	3.9	380	2 C83171	hypothetical prote
40	120	3.9	746	2 T83957	coatomer complex b
41	119.5	3.9	796	2 B85557	erythroonolide synt
42	119.5	3.9	3573	2 S23070	outer membrane por
43	119	3.9	856	2 T00349	Avicelase III - As
44	119	3.9	856	2 T00349	Avicelase III - As
45	118.5	3.8	631	2 T35234	probable secreted

ALIGNMENTS

RESULT 1

quinohemoprotein ethanol dehydrogenase (EC 1.1.99.-) type 1 precursor - Comamonas tes

C:Species: Comamonas testosteroni

C:Date: 08-May-1995 #sequence-revision 21-Jul-1995 #text-change 02-Jun-2000

C:Accession: S62366; S62373; S65908; S52317

R:Stoovogel, J.; Kraayveld, D.E.; van Sluis, C.A.; Jongejan, J.A.; de Vries, S.; Dui

Eur. J. Biochem. 235, 690-698, 1996

A:Title: Characterization of the gene encoding quinohemoprotein ethanol dehydrogenas

A:Reference number: S62366; PMID:96184549

A:Accession: S62366

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-708 <STOI>

A:Cross-references: EMBL:X81880; NID:g663195; PIDN:CA57464.1; PID:g663196

A:Accession: S62373

A:Status: preliminary

A:Molecule type: protein

A:Residues: 32-54 <STO2>

R:de Jong, G.A.H.; Geerlof, A.; Stoovogel, J.; Jongejan, J.A.; de Vries, S.; Duine,

Eur. J. Biochem. 230, 899-905, 1995

A:Title: Quinohemoprotein ethanol dehydrogenase from Comamonas testosteroni. Purific

A:Reference number: S65908; PMID:95324580

A:Accession: S65908

A:Molecule type: protein

A:Residues: 32-50, 'X', 52-54; 477-483, 'X', 485-490 <DEJ>

A:Experimental source: ATCC 15667

C:Genetics:

A:Gene: qhdh

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

C:Keywords: oxidoreductase; quinoprotein

F:1-31/Domain: signal sequence #status predicted <SIG>

F:32-708/Product: quinohemoprotein ethanol dehydrogenase type 1 #status experimental

Query Match 18.9% Score 583; DB 2; Length 708;

Best Local Similarity 27.9%; Pred. No. 1,3e-32;

Matches 173; Conservative 95; Mismatches 271; Indels 82; Gaps 22;

3 PTSLIMASAGALALLAAPFAQVTPVT-----DELLANPP-AGEMISYGNQEN 50

12 PGRMWLLAACLG--SAAFAQTGPAQAQAAAVORVDGDFIRANAARTPMPPTIGVYAE 69

51 YHSPITQTTTENVQOLVWAKGQPK-VQVPLIHGVMYLANGVYQIAIDAKTGD 109

70 FYYSRIDQINAAVNDLILAWSYNLESTRGEATPVVDGIMVYASMSVYHAIPTGN 129

110 LMEHRROLPNATLNSFG--PTRGMALYGTNVYFVSMNHLVALDTATG-----QVTFD 163

130 KIWYTDPOIDRTGKGGCDVYVNRGVALMKGVYVGAMDGRLIALDAATKEVHWHTFE 189

164 VDRGGEEDMVSNSGPVYANGVYAGST-CQYSPGCGVSGHDSATGEELMRNFIP--- 219

```

Db 190 GOKG-----LITGAPRFKGVIIIGKRGAEYVRG-YITAYDAEFERKMRNFSVPGDP 244
Oy 220 -----RAGEBDETWGNDYFARMITGA-----WGQITTPDVTNLVHGSTAVGPASEIQ 268
Db 245 SKPEDESKRAARTW--DPSGKMWKAGGGGTMDMSMTFPAELMTMYVGTGNGSPWMSHKV 302
Oy 269 RCTPGGTLTGNTREAVRPDGTGEIWRHQTLPDRNMDQECTFEKMMVNVVQSTEMEGIL 328
Db 303 RSPKGDNLVILASIVALDPDGTGKMYHQTPEFGNMDYTSTQPMITLADIKI----- 353
Oy 329 OSINPNATGERRVLTGVCPTGTMMQFDATGEFELWARDNTYONMIESIDENG-IVTVA 387
Db 354 -----AGKPRKVIYLNAP-KNGFEFVLDRTNGKFIKAFNEVPVNMASGDKNGKIGI- 404
Oy 388 EDALKEIDVEYDVCPTFLGGRDWPSSAALNPDGSIYFIPLANNCYDMAVDO-EFT----- 442
Db 405 --AARODSKRQDAPVPGFYGAHNMHPSENPQGTGLVYLPAQNVNVLMDKKMEFNQAGP 462
Oy 443 ----SMDYNTSNVTKL-PRGKDMIGRIDAIDISTGRILMSVERAANYSPVLSTGGVYL 497
Db 463 GKPGSGTGMNTAKFPFNAEPKSKFGRLLAMDVPAOKAANSVHVSHPMNGGILTTAGNV 522
Oy 498 FNGGTDYFRALSOETGETLMOTRLATVAGQALISTEVDGMOYVAILA-GGVSYSGSLNS 556
Db 523 FQGTADRLVAYHAATEKLEMAPTGTGVVAAPSTYVMDGROYVAVGCMGVYG----- 577
Oy 557 ALAGERVDTAIGNAVYVVAL 577
Db 578 -LAARATERGPG-TYTFYV 596

```

```

RESULT 2
S14270
alcohol dehydrogenase, membrane-bound (EC 1.1.-.-) 72k chain precursor - Acetobacter po
C:Species: Acetobacter polyoxogenes
A:Variety: strain NB11028
C:Date: 21-Nov-1993 #sequence_revision 26-Jul-1996 #text_change 20-Jun-2000
R:Accession: S14270
R:Jamaki, T.; Fukaya, M.; Takemura, H.; Tayama, K.; Okumura, H.; Kawamura, Y.; Nishiyama
Biochim. Biophys. Acta 1088, 292-300, 1991
A:Title: Cloning and sequencing of the gene cluster encoding two subunits of membrane-bc
A:Reference number: S14270; MUID:91159482
A:Accession: S14270
A:Molecule type: DNA
A:Residues: 1-738 <TAM>
A:Cross-references: GB:D00635; NID:9216185; PIDN:BA00528.1; PID:g216186
A:Experimental source: strain NB11028
C:Complex: heterodimer of 72k and 44k chains
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: alcohol metabolism; membrane protein; oxidoreductase
F:1-35/Domains: signal sequence #status predicted <SIG>
F:36-738/Product: alcohol dehydrogenase 72k chain #status predicted <MAM>

```

```

Query Match 18.3%; Score 564.5; DB 2; Length 738;
Best Local Similarity 28.0%; Pred. No. 2.6e-31;
Matches 165; Conservative 85; Mismatches 261; Indels 79; Gaps 14;

Oy 10 SAGALALLAPAFQVTPVDE-----LLANPPAGEWISYGCNENYHSPILQITTT 61
Db 17 TAGTICALLISYATMASADQAGTGAIIHADHPENMMTYGRTYSQDITSLDQINR 76
Oy 62 ENVGQQLVWAMGMPGVQV-TPLIHDGVWYLANPGDVIOAIDAKTGDLTWEHRRLP- 119
Db 77 SVNGMLKLAWYIDLDJNNGOECTPLVDGVYATTNMSMKAVDAATGKILMSYDPRPG 136
Oy 120 NIATINSEPTRGALGTWYFVSMNHLVALDTANGVTEFVNDGQEDAMSN----- 175
Db 137 NIADGGCCDYVNRGAAYMGKVFETGRLIADAKTGKILMSYVNTTIPPAELGKORSY 196
Oy 176 --SSGPIYANGVIAGSTCOYSPFGC--FVSGHDSATGEELMRNYFPRAGEED----- 226

```

```

Db 197 TVDGAAPTAKGVIIIGN--GGSEFGANGVSAFDAETGVDMREFTVNPKNEDPAASDS 254
Oy 227 -----ETMGNDYFARMITGA-----WGQITTPDVTNLVHGSTAVGPASEIQ 269
Db 255 VLMKNAYQTS-----PTGAMTRQGGGTWDSIYDPAALVYLIGNGSPWMYKTR 307
Oy 270 GTPGGTLTGNTREAVRPDGTGEIWRHQTLPDRNMDQECTFEKMMVNVVQSTEMEGIL 329
Db 308 SEGKQDNLFLSISYALKPBTGEYVWHFQETPMQDMFTSDQIMTLDLPL----- 357
Oy 330 SINPNATGERRVLTGVCPTGTMMQFDATGEFELWARDNTYONMIESIDENG-IVTVA 387
Db 358 -----NGETRHVYIARKNMGFFYIITDAKTGEISGKNVYVNMASGLDPTGRPIYP 410
Oy 389 DALIKEIDVEYDVCPTFLGGRDWPSSAALNPDGSIYFIPLANNCYDMAVDOEFT-SMDY 447
Db 411 DALYTLGKEWYIGPGLGHNFAAMFSPKGTLYIIPAOQVPELTNGVGFTHHPDSW 470
Oy 448 NTS---NVTKLPRG-----KDMIGRIDAIDISTGRILMSVERAANYSPVLSTGGV 496
Db 471 NLGIDMKKVGIPSPSEAKQAFVVDLKGWIVAMPQKQAEAMRVDRHGPWNGGILATGDL 530
Oy 497 LFNGGTDYFRALSOETGETLMOTRLATVAGQALISTEVDGMOYVAILAAG 546
Db 531 LFOGLANGEFHADVATNGSDFHPAADSGTIAPPYTILANGQYAVVEYG 580

```

```

RESULT 3
A49340
alcohol dehydrogenase (EC 1.1.-.-) precursor - Acetobacter pasteurianus (strain NC113
C:Species: Acetobacter pasteurianus
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 20-Jun-2000
R:Accession: A49340
R:Takemura, H.; Kondo, K.; Horinouchi, S.; Beppu, T.
J. Bacteriol. 175, 6857-6866, 1993
A:Title: Induction by ethanol of alcohol dehydrogenase activity in Acetobacter pasteu
A:Reference number: A49340; MUID:94042848
A:Accession: A49340
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-742 <TAM>
A:Cross-references: GB:D13893; NID:9517067; PIDN:BA04052.1; PID:g452586
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: alcohol metabolism; NAD; oxidoreductase

```

```

Query Match 18.2%; Score 561; DB 2; Length 742;
Best Local Similarity 28.5%; Pred. No. 4.6e-31;
Matches 173; Conservative 90; Mismatches 248; Indels 96; Gaps 21;

Oy 11 AGALALLAPAFQVTPVDELL--ANPPAGEWISYGCNENYHSPILQITTTNNGQLQ 68
Db 23 AALPLPAVPAADQAGTGEAIIHADHPENMWSYGTSTSEQRYSPILDINSWGDLEK 82
Oy 69 LVWARGMPGVQV-VTPLIHDGVWYLANPGDVIOAIDAKTGDLTWEHRRLP-NIATLNS 126
Db 83 LAMWYTLDTNRCQEAETPLVDGIMVATTNMSKMEALDAATGKILMQDPRVPGNIADKGC 142
Oy 127 REEPTRGALVITNYEYFVSMNHLVALDTANGVTEFVNDGQEDAMSNSSGPT- 180
Db 143 CDYVNRGAGYMGKVFQWCTFGDGRVLAADAKTGKRVMEVNTIPADASIGKORSYVDAVAV 202
Oy 181 VANGVIAGSTCOYSPFGC--FVSGHDSATGEELMRNYFPRAGEED----- 226
Db 203 VAKGLVILGN--GGSEFGARGFVAPDAETGKILMRVYTPNNKNEBDAVADNVLMSKA 260
Oy 227 -ETMGNDYFARMITGA-----WGQITTPDVTNLVHGSTAVGPASEIQRTPGTLY- 277
Db 261 YKTWGP--KGAAMVROGGGTVMDSLVYDPSDLTY--LAVG-----NSPPNNYRYRBE 309
Oy 278 --GTN-----TFFAVRPDGTGEIWRHQTLPDRNMDQECTFEKMMVNVVQSTEMEGLOSI 331
Db 310 GIGSMPLIGSIVALKPBTGEYVWHFQATPMQDMFTSDQIMTLDMPV----- 357

```



```

Db 357 LVLFYKKAKDKIVKATAHADPNNGFFIYVDNSNGKLNQANFPFVDNTTASHITDKTG--- 413
QY 365 WARDNTYONMIESIDENGIVTWNEDAILKELDEVYDCPTLGRDWPSSALNPDGSIYE 424
Db 414 --RPVERGSGORPLPEPG-----QKHGKAVEGSPPLPGGKNMMPAYSDDTGLIFY 461
QY 425 IPLNVGCDMAVADDEFMSMDVYNTSNTKLPCKDMIGRIDALIDISTGRTLSVERAAA 484
Db 462 VPANMKEDYITEEVSYSYKSGAYLGMGRFRIRMYDDHVSGLRADPVGKAVYWEHKEHLP 521
QY 485 NYSPLSTGGVLENGTDRYFRALSOETGETLMTQTRLATYVAGSAISYEVDGMOYVIA 544
Db 522 LMAGVLATAGNLVFTGTGDGYEFKAFDAKSKGELMKFQGSISIVPPLTWEDGEGQYIGVT 581
QY 545 GGGVSYGSGLSALAGEVNDST---AIGNAVYVYVFLP 578
Db 582 ---VGYG-GAVPLMGDMADLTREPVAOGGSPFWFKLP 614

```

RESULT 6

```

J00706
alcohol dehydrogenase (acceptor) (EC 1.1.99.8) alpha chain precursor - Methyllobacterium
N:Alternate names: methanol dehydrogenase 62K large chain
C:Species: Methyllobacterium extorquens
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Oct-1999
C:Accession: J00706; S07908
R:Anderson, D.J.; Morris, C.J.; Nunn, D.N.; Anthony, C.; Lidstrom, M.E.
Gene 90, 173-176, 1990
A:Title: Nucleotide sequence of the Methyllobacterium extorquens AM1 moxP and moxJ genes
A:Reference number: J00706; M0ID:90337342
A:Accession: J00706
A:Molecule type: DNA
A:Status: nucleic acid sequence not shown; translation not shown
A:Residues: 1-626 <AND>
A:Cross-references: GB:M31108; NID:g150017; PIDN:AAA25380.1; PID:g150018
A:Experimental source: strain AM1
R:Nunn, D.N.; Day, D.; Anthony, C.
Biochem. J. 260, 857-862, 1989
A:Title: The second subunit of methanol dehydrogenase of Methyllobacterium extorquens AM1
A:Reference number: S04644; M0ID:89350892
A:Accession: S07908
A:Molecule type: protein
A:Residues: 28-50, 'XX', '53' <NUN>
A:Note: the source is designated as Methyllobacterium extorquens AM1
C:Comment: This enzyme oxidizes methanol to formaldehyde.
C:Genetics: moxP
A:Keywords: alcohol metabolism; oxidoreductase
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-626/Product: alcohol dehydrogenase (acceptor) #status experimental <MAT>

```

Query Match 15.2% Score 468.5; DB 2; Length 626;
Best Local Similarity 25.4%; Pred. No. 9.3e-25;

Matches 159; Conservative 113; Mismatches 271; Indels 83; Gaps 23;

```

QY 10 SAGATALLA-AP-AFAQVPTVDLILANPAGE-WISYQONENYHSPLOITITENWGO 66
Db 7 SVSALMLALAPALSSGAYANDKVELSKSDNNWPKRNTSDNNFSLKQINKGNKQ 66
QY 67 LQLVW--ANGMOGKQVYPLIHGQVMT--ANPGDVIALDAKTGDLIWEHR-RQLPNI 121
Db 67 LRPAMWFTSTGLNGH-EGAPLVVDGKMYIHTSPNNTFALGLDDPGTILMOQKPKQNPA 125
QY 122 ATLSNGEPTKCALYGTN-----VYFVSMNHLVALDTAGQVTFYVDRCQGDGMVSN 175
Db 126 RAVACCDLVNRLGIATYPCGKTPALLKTLQDGNVAALAEETVWKEVNSDIDVGSSTL 185
QY 176 SSGPIYANGVIVAGSTQCYSPFCFVSGHDSATGEELMNFYI-----P 219
Db 186 TIAPVYVKDKVILIGSSGALGVRLTAVDVKTGEQVWRAVATGDKDLLASDPNINMP 245

```

```

QY 220 RAGEE--GDETWGNDYEARMWMTGA---WGQITVDPVTLVHYGSTAVGASSETQGTGPG 274
Db 246 HYCGKIGTGTWEGD---AMKIGGTMWGYAYDPCTNLITFOTGNPAPWNETMR--PBD 300
QY 275 TLVTNTFRFAVRPDGTETIWRHOTLRDNDQDCTEEMAMTVNDVOPSTYEMGLOSINP 334
Db 301 NKW-TWTFEGRADTGEAKGYOKTPHDEWD-----YAGVNVAMLSKORD----- 344
QY 335 AATGERFVLTVGVPCKTGTMMQFPAETGEFLMAR---DTYVONMIESIDENGIVTWNEDAI 391
Db 345 -KDGKARKLLTHDRNGIYITLDRTGALVSANKLDDT--VAVFKSVDLKTQOPRPADPEY 401
QY 392 LKELD-VEVDVCEPTLIGRDPSPALNPDGSIYEIPLNVGYDMAVADDEFMSMDVYNTS 450
Db 402 GTMDHLAKDICPSAMGYHNOGHSYDPRRELFEMGINICHMDWEPFLPYRAGQFVYGA 461
QY 451 NVTKLPPGK-----DMIRIDAIDISTGRTILMSVERAANSPVLSTGGVLFNGGDR 504
Db 462 TLNMYPGPKDQKQNYEGDQIKAYNAITGDIYKMKERFAVWGVTAAAGDLVFFYGLDG 521
QY 505 YFRALSOETGETLMTQTRLATYVAGSAISYEVDGMOYVAI---AGGGSYSG-----S 552
Db 522 YLKARSDTGDLIMKFKIPSGAIGYPMYTHKGTOYVATYVGGMPCGLVFDLADPTA 581
QY 553 GLNSALAGERV-DSTAIIGNAVYVFLP 577
Db 582 GLGAVGAFKRLNNTYTMGGGVVYFSL 607

```

RESULT 7

```

S68591
methanol dehydrogenase (EC 1.1.1.244) heavy chain - Methylotrophus methylotrophus (str
C:Species: Methylotrophus methylotrophus
A:Variety: strain W3A1
C:Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 08-Oct-1999
C:Accession: S68591
R:Xia, Z.; Dai, W.; Zhang, Y.; White, S.A.; Boyd, G.D.; Mathews, F.S.
J. Mol. Biol. 259, 480-501, 1996
A:Title: Determination of the gene sequence and the three-dimensional structure at 2.
A:Reference number: S68591; M0ID:96256524
A:Accession: S68591
A:Status: Preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-573 <XIA>
A:Cross-references: EMBL:U04040; NID:g1127819; PIDN:AAA83765.1; PID:g1127820
A:Note: the authors did not translate the codons for residues 1-2
C:Keywords: oxidoreductase

```

Query Match 14.2% Score 439.5; DB 2; Length 573;

Best Local Similarity 24.5%; Pred. No. 8.4e-23;

Matches 146; Conservative 99; Mismatches 259; Indels 93; Gaps 21;

```

QY 38 AGEWISYQONENYHSPLOITITENWGOQLVW--ARGMOGKQVYPLIHGQVMTL-- 93
Db 13 AGAMPVATGYYTQSHNSPLAQINKSNVKNVKAWSFTGVLNGH-EGAPLVIGDMNVVHS 71
QY 94 ANPGDVIALDAKTGDLIWEHR-RQLPNIATLNSFGEPTKCALYGTNYYFVSMNHLVA 152
Db 72 AEPNNTYALNLDNGKTIWQHPRKODASTRAVMCCDVADVAGLAVGAGQIYKQANGHLA 131
QY 153 LDTAGTQVTFPDNDGQGDGMVSNSSGPIYANGVIVAGSTQCYSPFC--FVSGHDSATGE 210
Db 132 LDKATGKINWEVCECDPKVSGSTLQAPFVAKDVLWG--CSGAEGLGVAGVAVNAPDLKTGE 189
QY 211 ELWR-----NYFIPRAGE--EGDETWGNDYEARMWMTGA---WGQITVDP 249
Db 190 LKWRAPATGSDSVYRLAKDFNSANPHYGQGLGKTKWEGD---AMKIGGTTMNGWAIYDP 246
QY 250 VTNLVHYGSTAVGASSETQGTGPG---TLVTNTFRFAVRPDGTETIWRHOTLRDND 304
Db 247 KLNLFYSGNPAWNETMR--PBDNKWTMTWIGRDL-----DTGAKWGYOKTPHDEW 298

```


Db 593 -----VPPGMLGVFENGCMGVNPDROVAVANNPITGLPFVSRILIPADPNRAQTAKG 641

Qy 426 -----PLNNVCYDMAVDEFTSMDEVNTSNVTKLPKPGKMDIGRIDADISTGRILM 477

Db 642 AGTEGCVQPMGVVY-----GVEISAFSLPLGLPCPKPAMGVYAGVDLKTHEVW 691

Qy 478 -----SVERAANYSPLVSTGGVLTNGST-DREYFALSOEGETIM 518

Db 692 KKRIGTIRDSLPNLFQLPVAKIGVPLGSLISTAGVNVFVAGQDNYLRAFVNTGKALM 751

Qy 519 QTRLATVASGA--ISYEVDGMQYVAI-AGGVSYSGSLNSALAGERVDSIAGNVYVE 575

Db 752 EALLP--AGGQATPMTEYINGKQYVIMAGHSFC-----TKMGDYLAVY 795

Qy 576 ALP 578

Db 796 ALP 798

RESULT 11

QPKEX

glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - Glucobacter oxydans

C:Species: Glucobacter oxydans

C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 22-May-1998

C:Accession: S17716; S19265

R:Cleton-Jansen, A.M., Dekker, S.; van de Putte, P.; Goosen, N.

Mol. Gen. Genet. 229, 206-212, 1991

A:Title: A single amino acid substitution changes the substrate specificity of quinoprot

A:Reference number: S17716; MUID:92017653

A:Accession: S17716

A:Molecule type: DNA

A:Residues: 1-808 <CLE>

A:Cross-references: EMBL:X62710

R:Goosen, N.

submitted to the EMBL Data Library, February 1992

A:Reference number: S19265

A:Accession: S19265

A:Molecule type: DNA

A:Residues: 1-212, 'A', 214-808 <GOO>

A:Cross-references: EMBL:X62710; NID:958416; PID:958417

C:Genetics:

A:Gene: gdh

C:Function:

A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone

A:Pathway: respiratory chain

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

C:Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembrane

F:9-28/Domain: transmembrane #status predicted <TM1>

F:35-54/Domain: transmembrane #status predicted <TM2>

F:60-76/Domain: transmembrane #status predicted <TM3>

F:94-110/Domain: transmembrane #status predicted <TM4>

F:122-138/Domain: transmembrane #status predicted <TM5>

F:91/93/Binding site: ubiquinone (Air, Asp) #status predicted

F:470/Active site: Asp #status predicted

Query Match 11.0%; Score 340; DB 1; Length 808;

Best Local Similarity 21.8%; Pred. No. 1,1e-15;

Matches 159; Conservative 98; Mismatches 259; Indels 212; Gaps 30;

Qy 13 ALALLAFAFAQVPTVDLLA-----NPPAGEWISYSGONENRHSPLTQIT 61

Db 128 AVLALFASLFDPPDISGLPTQIANASPADPNDVPASEMHAYGRTOAGRWSPINQINA 187

Qy 62 ENVGQDLVW-----ARKQOPKV--QVTPLLHDGVMTANGVDYIOAIDATGSLIW 112

Db 188 TNSVILKAVAMHITKDMNNSNDPEQINEATPIEFNNLTLYCSLHQKLEAVDAGATGVAKW 247

Qy 113 EHRRL-----PNLATLN-----SFGG-PTRGMALYGT-----NVYFVSDMHLVALDT 155

Db 248 YDPRKLTQINPFGHILTCGVSFHEPRANAMSDGNPAFTDCAKXSLIPVN-DGRLEVEDA 306

Qy 156 ATGVTFDV-DRGGEDMVSN-----SSGPIYANGVIVA-----GSTCQYSP 196

Db 307 DTGKTCGFGNNGEIDLVRPNQPTTFBQYEPSTPVTDKLIIINSAITDNGSVKQASG 366

Qy 197 FCGEYSGSDASTGELW-----RNFIFRAGEGDETNENDVFAWMGTGAMQIYDPTN 252

Db 367 --ATQAFDYVTGRKRWVFDASNDPNQULPDSEHPVFRPNPSNVIYS-----SYDANLN 418

Qy 253 LVHGSTAVGPASETQKPT--GGLYGTNTRF-----AVRPDTGELVWRHQTLPDNN 304

Db 419 LVYIPMGV-----CTPDQMGDRTKDSERFAPGIVALNADTGLKAMPYQVHHDLW 469

Qy 305 DQCTFEEMVTVNDVQSTEMEGLQSLNPNAATGERRLTVCPKTKTMMQFDETG-EF 363

Db 470 DMELPSQPSLVDVQKDGTLVPAIYA-----PKTGDIFVLDRTGKEI 513

Qy 364 LMARDT-----NYQNMIESIDENGIVTVNEDALIKELDY----- 397

Db 514 VPAPETPVQGAAPGDHNSPTQPMSSQ--LTLRPKNPLNDSIDINGGTFIDMPCSIYFHT 570

Qy 398 -EYD-----VCPTEFGGRDMPASALNPDGTYIFPLNNVCYDMAVADQ----- 439

Db 571 LREBPTTPPSLAKSLIRPGDLGMEWGLAVDPQROVAFANPISLPSVQLVFRGPNP 630

Qy 440 -----EFTSMDEVNTSNVTKLPKPKDM-----GRIDAIDIST 472

Db 631 IMPERNAKGTGETGLQHNHYGIPYAVNLPFLDVLPLFPGIKMPCRPWGVAGIDILKT 690

Qy 473 GRTLSVERAANYS-----PV-----LSTGGGLVF-NGSTDYFRALS 510

Db 691 NKVVMOHNRNGTIRDSMYGSSSLPPLPIKIGVPSLGGSLTAGNIGFLTASMDYITRAYN 750

Qy 511 QETGETLMQTRLATVASQALSYEVDGMQYVAIAGGVSYSGSLNSALAGERVDSIAIGN 570

Db 751 LTTGKVLWMDRLPAGQATPTTYAINGKQYI-----VTYAGGHSNF-----PTRMGD 797

Qy 571 AVYVFLAP 578

Db 798 DIIAYALP 805

RESULT 12

F83360

glucose dehydrogenase PA2290 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: F83360

R:Stoeyer, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Adam, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa

A:Reference number: A82950; MUID:20437337

A:Accession: F83360

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-803 <STO>

A:Cross-references: GB:AE004654; GB:AE004091; NID:99948311; PID:AA605678.1; GSPDB:GN

A:Experimental source: strain PA01

C:Genetics:

A:Gene: gcd; PA2290

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 10.6%; Score 327; DB 2; Length 803;

Best Local Similarity 21.9%; Pred. No. 8,6e-15;

Matches 166; Conservative 84; Mismatches 265; Indels 244; Gaps 32;

Qy 6 LLM-----ASAGLALLAARPAFAQVPTVDLLANP----- 36

Db 103 LLMVPMFRRLADGAPRLGTALGVAIVLAGAAAGSO-FTNPGQIVGRIDRDSGMTST 161

Qy 37 -----PAGEWISYSGONENRHSPLTQITTEVNGQLDLYVARGMQPQKV-----Q 81

Db 162 APAMPDGMQAYGRTFEGRDYSPLKQITPANVGQLEAMR--IRTDGLPTADPLELTNE 219
 QY 82 VTPLIHGVWYLANPGVIOAIDAKTGDLIWEHRQPLNATLSPFEP--RGALYGTN 140
 Db 220 NTPFLKVMGLACTAHSAHSAVLAADPDGAEIWRPDQIOSPVGKGFAMHMCRCGVSYDEE 279
 QY 141 VY-----FVSWDNLVALDTATGQV-----TFDV 164
 Db 280 QYASDVCAFPALSEAGKAVAASCPRRLFLPTADARLAIINANGKVCDFGKGAVDL 339
 QY 165 DRGGE---DMVNSSGPIYANGVYAG---STCOYSPGCFVSGHDSATGEIMRNY 216
 Db 340 TAGIGPTPGGYSTSPAAYTRNLVIGIIVHTDNESSTEPSG--VIRAFVDHDKLVW--NM 397
 QY 217 FIPRAGEGDETWGNDYEARMGTGAMGQITPDYPTNLVHGSTAVGASSTGRC---TPG 273
 Db 398 DSGNPDETEPLAPKFY--TRNSPMMMSLASVDEKLGQY--LPLGQMPDQMGKNTPG 453
 QY 274 GTLYGTNTRFAVRPDTEIYWRHQTLPDWMDOCTEFEMVTVNDVOPSTEMEGLSINP 333
 Db 454 AEKFSAGL--VALDINTGKLRNMYOFTHDLMDM-----VGSQPTL-----LDL 496
 QY 334 NAATGEREVLTVGRCRTGMQPDATG-----361
 Db 497 KTADGVAPALI--APTKGSLYVLDRRDTPVPIREVPAGQAVEGDHTAPQARSDNL 555
 QY 362 -----EPLMA-----RDTNYONMIESIDENGIVTMDATLKELDEYDY 401
 Db 556 LRPLTERDMWSSPFDQMCRIQFRSLRYEGQYTPPSEGSGL-----I 599
 QY 402 CPPTFLGRDMPASALNDPSGIYFPLNNVCYDMAVDOEFT---SMDVNTSNVTK----- 454
 Db 600 YPGNVGVFNMGVSVDPDRQILFTSPNYMAFVQMWPRDKVPGSKREGTSCVQNTGA 659
 QY 455 -----LPGKMGIRIDAIDISGRTILMSVERAA--NYSY----- 489
 Db 660 PYAVIMHPESPISGLPQAPSWCDVAGIDLTAKVWQHNGTSRNTPTPIGLTVGVPS 719
 QY 490 -----LSTGGVLENGGT--DRYFRALSOETGELTMQTRLATVASQA--ISYE--VDGMOYV 541
 Db 720 MGGSITTAGVAFPLSGTLQDLARAYVKDQKQKOLMARLP--AGQAATPMSYTKDKGRQYV 777
 QY 542 AI--AGGCVSTGSGUNSLAGERVDSALGNAYVYFALPQ 579
 Db 778 LIVAGGHGSPG-----TRMGDIYIYALPR 802

RESULT 13

A55547
 quinase-shikimate dehydrogenase (EC 1.1.99.-) - Acinetobacter calcoaceticus
 C:Species: Acinetobacter calcoaceticus
 C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
 C:Accession: A55547
 R:Elasmere, D.A.; Ornston, L.N.
 J: Bacteriol. 176, 7659-7666, 1994
 A:Title: The pca-pdb supraoperonic cluster of Acinetobacter calcoaceticus contains quiaA
 A:Reference number: A55547; MUID:95095936
 A:Accession: A55547
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-809 <EIS>
 A:Cross-references: GB:U11554; NID:9508824; PID:9508825
 C:Genetics:
 A:Gene: quia
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
 C:Keywords: membrane protein; oxidoreductase

Query Match 10.1%; Score 313.5; DB 2; Length 809;
 Best Local Similarity 23.3%; Pred. No. 7.5e-14;
 Matches 160; Conservative 80; Mismatches 262; Indels 185; Gaps 26;
 QY 3 PPSLWASGALALALLAA--PAFAQVTPVDELLANPAGEMISYGNQENYRHSPLTQITTT 61

Db 148 PHETVKAAGEELPLVEVDPAKKOV-----NWDHYGNDAGSRFVALDQINR 193
 QY 62 ENVGQOLVW-----ARGMQPG--KVQVTPLIHGVWYLANPGVIOAIDAKTGDLIWE 113
 Db 194 NNYSKILKEAMRFRTGDEFTGTGNGAEDQMTPLQVGNKAVFLCPHNNIFRAIDDSQKQMLK 253
 QY 114 HRQPLNATLNSFG--EPTRGMALYGT-----NYFV 144
 Db 254 -----AEVNSTADAMERCGRVAFYFSTOPLVOPPLAGATPVAALANTECPRRVYTN 305
 QY 145 SWDNLVALDTATG-----QVFEDVDGQGEDMVS-----SSGPIYANGVYAGSTC 192
 Db 306 TYDGRILAVNADTGACRCFGVNGVYNLHEGIGENTKAPREFVTAAPITAGTIYVGSRI 365
 QY 193 QYSPF---GCFVSGHDSATGEELMRNYFIPRAGEGDETWGNDYEARMGTGAMQIYTD 248
 Db 366 ADVNADMPGQVIRAVDYITTKLRMA--FDPNPNPNVYLKPGELIKRSSSTSMAMASGD 423
 QY 249 PVTNLYH--YGSTAVGASETQRTGTLGTNTRFAVRPDTEIYWRHQTLPDWMDO 306
 Db 424 PQMTVFLPMGSSSVDVWGNR--TAADHKYNTSV--LALDATTKGKQVYNTVHNDLMD- 479
 QY 307 ECTFEMVTVNDVOPSTEMEGLSINPNAATGEREVLTVGRCRTGMQPDATG----- 364
 Db 480 -----FDLPMQPS-----LVDFPMKDGITKRAVYIGT--KSQFVLDRTYGRKPLTK 524
 QY 365 -----WADTNYONMIESIDENGIVTNE-----DALIKELD--VEYD 400
 Db 525 VLEOPIKVADIPGEQYSTQPSVEMPOIGNOTLKESDMGATPPDQLMCRINFSKMYD 584
 QY 401 -----VCPPTFLGRDMPASALNDPSGIYF-----IPL 427
 Db 585 GLYTPAGTDVLSIFPGSLGGMWGSIAFDPTHRVYFVNDMRIGLITQIKOTPEDIKQA 644
 QY 428 N---NVCYDMAVDOEFTSMDVNTSNVTKL--PPGKDMIRIDAIDISTGRTILMSVERA 482
 Db 645 NGGEKVNNGMAVPMKGPYKKNRKFMSALGIPQCKPFTMAIDMKTKQVAMQVPLG 704
 QY 483 AANYT-----PVLTSGGVLFNGCTDRYFRALSOETGELTMQTRL 522
 Db 705 TIQDTPGMIKGLKAPIGMPTIGGPMATOGGLVFATQDYVLRAPVSSNGKELMARL 764
 QY 523 ATVASQAISY--EVDGMQVYAIAGGG 547
 Db 765 PYGSGQTPMSTMSPTKQYVVSAGG 791

RESULT 14

JC4881
 polyvinyl-alcohol dehydrogenase (acceptor) (EC 1.1.99.23) precursor - Pseudomonas sp.
 N:Alternate names: PVA dehydrogenase
 C:Species: Pseudomonas sp.
 C:Date: 10-Sep-1996 #sequence_revision 18-Oct-1996 #text_change 08-Oct-1999
 R:Shimao, M.; Tamogami, T.; Nishi, K.; Harayama, S.
 J: Bioest. Biotechnol. Biochem. 60, 1056-1062, 1996
 A:Title: Cloning and characterization of the gene encoding pyrroloquinoline quinone-d
 A:Reference number: JC4881; MUID:96376165
 A:Accession: JC4881
 A:Molecule type: DNA
 A:Residues: 1-639 <SHI>
 A:Cross-references: DDBJ:D50670; NID:91502278; PIDN:BA09321.1; PID:dt1009963; PID:915
 C:Comment: This enzyme is a membrane-bound enzyme. It oxidizes low molecular weight s
 C:Genetics:
 A:Gene: pdh
 C:Keywords: alcohol metabolism; oxidoreductase
 F:1-33/Domain: signal sequence #status predicted <SIG>
 F:49-53/Region: gene C-binding
 F:294-497/Domain: pyrroloquinoline quinone-binding #status predicted <POB>

Query Match 9.0%; Score 276.5; DB 2; Length 639;

Best Local Similarity 22.4%; Pred. No. 2e-11; Matches 143; Conservative 70; Mismatches 241; Indels 183; Gaps 28;

QY 7 LMSAGALLAAPAPAOQTPVTDLLANPAGEWISYQONENRHSPLDITTEENVOQ 66
 DB 117 IWPFPASMPDLPCKKKIPID--LSTP--DQMGWAGITNAFQNPGLTADVP 172
 QY 67 LQVMAKQOPKQVQVPLIHGVTYLANPBGVIAIDAQTDGLWEHROPLNATLNS 126
 DB 173 LKYMAFN-YPGSKNGQATVVGDRLEFVMSGAVYALNKTGCVWRHDAANAATSSVAV 231
 QY 127 FGEPTGMAALGTYNVEFVMDNHLVALDTATG---QVTFVDYRG-----QGEDM 172
 DB 232 VOLPAGAPAOYA--IFFSDMTAAVALDQOTKOLMKTIDQPGVQMTGSPTEHGLKF 289
 QY 173 VNSSGPIYANGVIYVGSFCOYSPGCEFGSHSATGEELMNYFLP-----RAGEED 226
 DB 290 VPISSG---NEAFATNDQMECCFRGALVALDALSGKVMKTYTQKEPAPFLNKLGK 345
 QY 227 ETWGNDEYARMWMTGA--WQIITYDPVTNLV-----HYSTAVGPASETQRTGP 272
 DB 346 QMWG-----PAGSITWGAPTIDPKRGLVYVATSNSTYEVHESDAV----- 387
 QY 273 GGLTYGNTFRPAVRPDTGTYVRHQTLPNDNDQCTFEMAVTVNDVOPSTMEGLQSTN 332
 DB 388 -----NAMEIETGVRMINOVTKDNYIIGCP---RAANCP-----EKYG 424
 QY 333 PNAATGERHVL-----TGVPCKTGTMOQDAET--GEFLARDTNYQNIESIDENG 382
 DB 425 PDFALGNSPILTLTDGROYIYVGOKSGAVYAMDNDGELTMR----- 469
 QY 383 IYTVNEDALIKELDEYDVCTFLGRDWPSSALNPDGTYFPLNNVCYDMMAYDQFT 442
 DB 470 -----RVSPGSEL---GGVEF-----GMAADAE-- 489
 QY 443 SMDVY-NTSNVTKLPKPKMIGRIDALIDISTRTLMS--VERAANY-----SPVL 490
 DB 490 --NVYVIGISDVITRRGKRP--GYVALIRIDGADYAWAPPAPRTPCRMNNIFCHPAVSQAV 544
 QY 491 STGGGVLFNGSTDRYFRALSOETGETLMOTRLA-----TVASGOAISYVDQMQVAVAG 545
 DB 545 TAMPGVYFAGSMGHRFASTSDGKVLMEFNRAAPRYKTIVAGKQADGGMQAG-PTIAG 603
 QY 546 GGVSYSGSINALAGERVDSTA-----IGNAVYFAL 577
 DB 604 GMVYVHSG---YAGRSTQNGADLRGREGNVLIAFSV 636

RESULT 15

A41378
 hypothetical protein (cyb 5' region) - Paracoccus denitrificans (fragment)
 C:Species: Paracoccus denitrificans
 C:Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 08-Oct-1999
 C:Accession: A41378
 R:Ras, J.; Reijnders, W.N.M.; Van Spanning, R.J.M.; Harms, N.; Oltmann, L.F.; Stouthamer, J. Bacteriol. 173, 6971-6979, 1991
 A:Title: Isolation, sequencing, and mutagenesis of the gene encoding cytochrome c-553i c
 A:Reference number: A41378; MUID:92041583
 A:Accession: A41378
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-221 <RAS>
 A:Cross-references: GB:M75583; NID:g150576; PIDN:AAA25574.1; PID:g150577

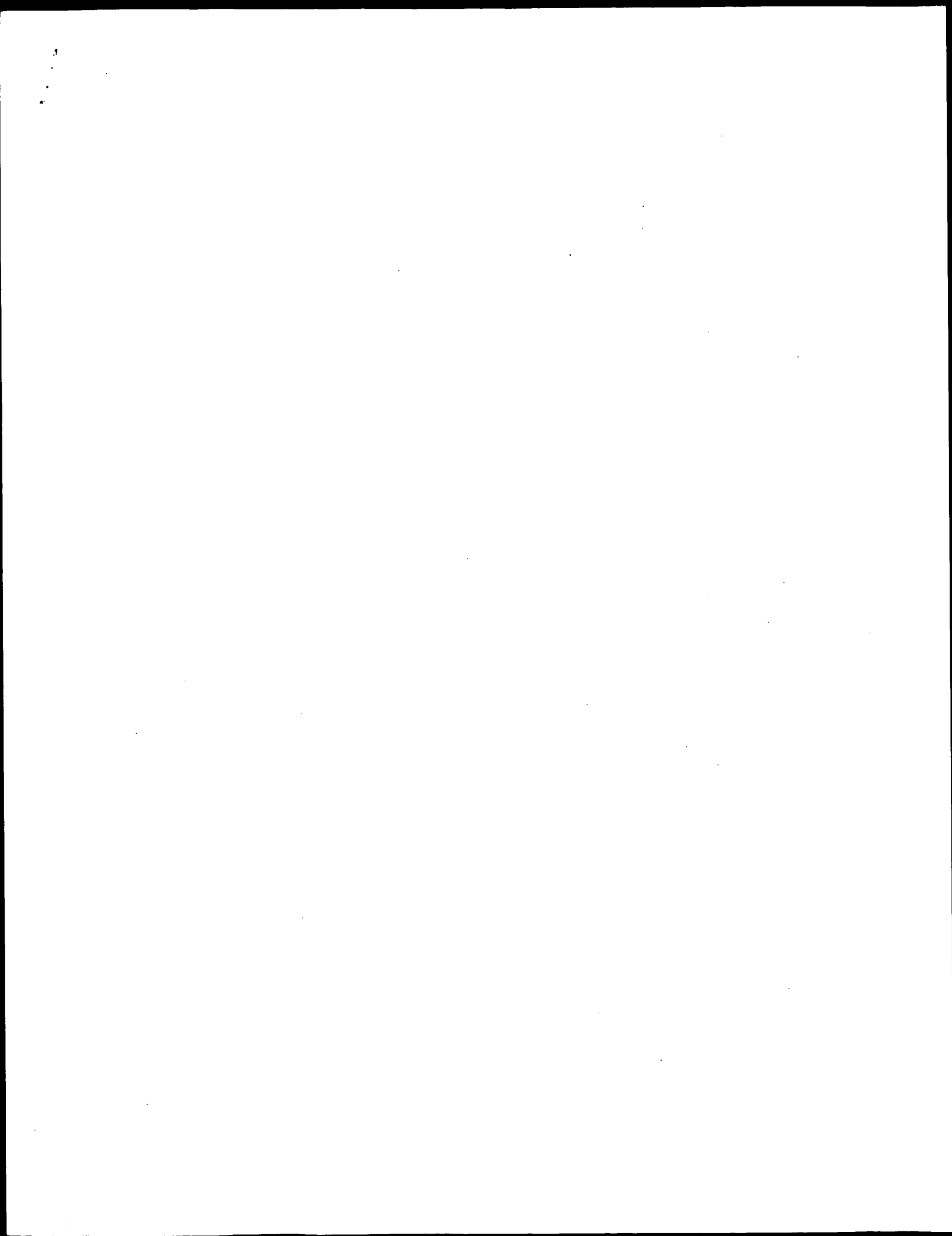
Query Match 7.2%; Score 221.5; DB 2; Length 221;

Best Local Similarity 30.1%; Pred. No. 2.8e-08; Matches 58; Conservative 34; Mismatches 86; Indels 15; Gaps 4;

QY 401 VCPTEFLGGRDMSALNPDGTYFIFPLNNVCYDMMAYDQFTSMQV--NTSNVTKLPPG 458
 DB 28 VCPALGTRDQPPALFSPKTNLFYPTNHVCMDEPFRVATYTAGQPIYGATLSMTAPNS 87

QY 459 KDMIGRIDALISTGRTLSVERAANYSPVLSTGGVLFNGSTDRYFRALSOETGETLW 518
 DB 88 HGGMGNFIAHMTTGEIKWSVPEQFSVWSGALATAGDVYFGTLEGLKRPVDAQTGELY 147
 QY 519 QTRLATVASGALISTEVDQMAYVIAAGS-----GVSYGSLNS-----ALAGERVDS-- 565
 DB 148 KFKTPSGITGNVMTYEHGKQYVIGILSGVGWAGIGLAAGLTNPDGIGAVGYASLSQY 207
 QY 566 TAIGNAVYFALP 578
 DB 208 TELGGQLTFEELP 220

Search completed: August 8, 2001, 19:40:33
 Job time: 30 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 2001, 19:41:08 ; Search time 13.25 seconds
(without alignments) updates/sec
1496.900 Million cell

Title: US-08-934-506A-5

Sequence: 1 MKPTSLMASGALALLAAP.....GERVDSALGNVYVFPQ 579

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	569	18.4	739	1 DHET_ACEU	044002 acetobacter
2	564.5	18.3	738	1 DHET_ACEPO	P28036 acetobacter
3	553.5	17.9	757	1 DHET_GUSU	005542 gluconobact
4	547.5	17.7	742	1 DHET_ACEAC	P18278 acetobacter
5	505	16.3	623	1 EXAA_PSEAE	092477 pseudomonas
6	495	16.0	631	1 DHM1_PARDE	P12293 paracoccus
7	487.5	15.8	600	1 XOXE_PARDE	P29968 paracoccus
8	477.5	15.5	626	1 DHM1_METEX	P15279 methylolact
9	468.5	15.2	626	1 DHM1_METEX	P16027 methylolact
10	436.5	14.1	571	1 DHM1_METME	P38539 methylolact
11	392.5	12.7	796	1 DHG_ECOLI	P15877 escherichia
12	390	12.6	801	1 DHGA_ACIQA	P05465 acetobact
13	340	11.0	808	1 DHG_GLUOX	P27175 gluconobact
14	313.5	10.1	809	1 QUIA_ACIQA	059086 acetobact
15	276	8.9	790	1 QUIA_XANCU	092478 xanthomonas
16	155.5	5.0	392	1 YFGL_ECOLI	P77774 escherichia
17	145	4.7	1645	1 OMPB_RICTY	P96389 r outer mem
18	128	4.1	1365	1 GTEP_STRDO	P29336 streptococc
19	122.5	4.0	644	1 XYND_CELFT	P5465 cellulomona
20	121.5	4.0	799	1 AFESK_STRCO	P54741 streptomyc
21	121.5	3.9	827	1 YZ23_METUA	060282 metanococc
22	119.5	3.9	796	1 COPP_SCHPO	042937 schizosacch
23	119	3.9	443	1 PORP_PSEAE	P32722 pseudomonas
24	116	3.8	746	1 PEPA_ECOLI	P05825 escherichia
25	115.5	3.7	833	1 YXAL_BACSU	P42111 bacillus su
26	115.5	3.7	827	1 CSG_HALVO	P25062 halobacteri
27	115	3.7	3567	1 ERYZ_SACER	003332 saccharopol
28	113.5	3.6	790	1 PLMN_PIG	P06867 sus scrofa
29	112.5	3.6	1026	1 VG37_BPT4	P03744 bacterioph
30	111	3.6	2003	1 YDBA_ECOLI	P33666 escherichia
31	110.5	3.6	894	1 FOX2_NEUCR	001373 neurospora
32	109.5	3.5	1012	1 FOLG_IBDOV	P27276 avian infec
33	109.5	3.5	1643	1 OMPB_RICPR	Q53020 r outer mem

34	109	3.5	593	1 SPG2_STRSP	P19909 streptococc
35	108	3.5	2265	1 FINC_BOVIN	P07589 bos taurus
36	106.5	3.4	537	1 TEE6_STRPY	P18481 streptococc
37	106.5	3.4	649	1 ACES_DROME	P07140 drosophila
38	106.5	3.4	784	1 OSTA_ECOLI	P1554 escherichia
39	106	3.4	750	1 BPPX_STRPN	P14677 streptococc
40	105.5	3.4	1256	1 FINC_CHICK	P11722 gallus gall
41	105	3.4	1256	1 ATL_STNAU	P52081 staphylococ
42	104.5	3.4	679	1 FLAY_CAUCR	P15346 caulobacter
43	104.5	3.4	781	1 COAT_PAVHB	P07299 human parvo
44	104	3.4	499	1 VIL1_HPV13	002273 human papil
45	104	3.4	762	1 SLAP_ACEKI	P22258 acetogenium

ALIGNMENTS

RESULT	ID	Query Match	Length	Description
1	DHET_ACEU	18.4	739	044002 acetobacter
2	DHET_ACEU	18.3	738	P28036 acetobacter
3	DHET_GUSU	17.9	757	005542 gluconobact
4	DHET_ACEAC	17.7	742	P18278 acetobacter
5	EXAA_PSEAE	16.3	623	092477 pseudomonas
6	DHM1_PARDE	16.0	631	P12293 paracoccus
7	XOXE_PARDE	15.8	600	P29968 paracoccus
8	DHM1_METEX	15.5	626	P15279 methylolact
9	DHM1_METEX	15.2	626	P16027 methylolact
10	DHM1_METME	14.1	571	P38539 methylolact
11	DHG_ECOLI	12.7	796	P15877 escherichia
12	DHGA_ACIQA	12.6	801	P05465 acetobact
13	DHG_GLUOX	11.0	808	P27175 gluconobact
14	QUIA_ACIQA	10.1	809	059086 acetobact
15	QUIA_XANCU	8.9	790	092478 xanthomonas
16	YFGL_ECOLI	5.0	392	P77774 escherichia
17	OMPB_RICTY	4.7	1645	P96389 r outer mem
18	GTEP_STRDO	4.1	1365	P29336 streptococc
19	XYND_CELFT	4.0	644	P5465 cellulomona
20	AFESK_STRCO	4.0	799	P54741 streptomyc
21	YZ23_METUA	3.9	827	060282 metanococc
22	COPP_SCHPO	3.9	796	042937 schizosacch
23	PORP_PSEAE	3.9	443	P32722 pseudomonas
24	PEPA_ECOLI	3.8	746	P05825 escherichia
25	YXAL_BACSU	3.7	833	P42111 bacillus su
26	CSG_HALVO	3.7	827	P25062 halobacteri
27	ERYZ_SACER	3.7	3567	003332 saccharopol
28	PLMN_PIG	3.6	790	P06867 sus scrofa
29	VG37_BPT4	3.6	1026	P03744 bacterioph
30	YDBA_ECOLI	3.6	2003	P33666 escherichia
31	FOX2_NEUCR	3.6	894	001373 neurospora
32	FOLG_IBDOV	3.5	1012	P27276 avian infec
33	OMPB_RICPR	3.5	1643	Q53020 r outer mem

Query Match: 18.4% Score 569; DB 1; Length 739;
Best Local Similarity: 28.3%; Pred. No. 2.8e-32;

Matches 165; Conservative 85; Mismatches 268; Indels 66; Gaps 14;

QY 10 SAGALALLAARFAQVTVTDE-----LLANPAGEWISYCONENRHSPLTQIT 61
 Db 17 TAGTICALISGYATWASADGQATGEALIHADHPNMWMTYGRYSEGRYSPLDQINR 76
 QY 62 ENVGQLOLVWANGMOPGKQVY-TPLIHGVWYLANPGVIOAIDAKGDLIWEHRRLP 119
 Db 77 SNVGNLKLAWYLDLTNRGQEGTPLYIDGVYATTTNMSMAKAVAAIGKILMSTDPYRG 136
 QY 120 NIATLNSFGEPTRGMAIXGTNYEVSMDNHLVALDTATGQVTPVDKRGQEDVNSN- 175
 Db 137 NIADKGCDDTYNRGAAYWNGKVFYETFGRLIALDAKTGLWVSNTLPPAEELGKGRSY 196
 QY 176 --SSGPIVANGVIAGSTCOYSPGCG--FVSGHDSATGEELMRNYFLPRAGEEDTWCN 231
 Db 197 TYDGAPRIAKRVIIGN--GGSEFGARGVAFPAETGKVDWRFETAPNPKNEPDHNASD 254
 QY 232 D-----YEARMMTGA-----WGLTYDPVTNLVHGSTAVGPAPESTGRTGCT 275
 Db 255 SVLMNKAVQWSPGTGAMTRGGGGTVDMSIYDVPADLVYLGVNGSPWNYKYRSEKGD 314
 QY 276 LYGNTRAVAPDPTGEIYWRHQLTPRNDWDECTFEKMAVTNVQSTEMEGLOSLNPA 335
 Db 315 NLFGLSIVALKPTEGEYVWHQETPRMDQFTSVQOITMLDLPINGET----- 362
 QY 336 ATGERRVLTGVPCKGTGTMQOPDAETGEFLMARDNTYONMIESID-ENGIVTWNEDATLKE 394
 Db 363 -----RHVIYVAP-KNGFFIIDAKTGEFISGKIVYVWMAAGDLPKGRPIYNDALYTL 417
 QY 395 LDVEYDVCPTFLGGRDPSAALNPSGSIYFPLNNVCYDMAAVDOEFT-SMDVNTS--- 450
 Db 418 TGRMYGIPGDLGGHNPAAAFSPKGTGLYIIPAQVPEFLTNGVGTTPRDPDSMNLGLDM 477
 QY 451 NVTKLPPG-----KDMIGRIDALISTGRTLMSVERAANYSPVLSTGGCVLENGGT 502
 Db 478 NKVGIPSPSPKQAFVVDLKGWIVAMPDOKQAEAMRVDRHGRPNWNGILLATGDLFLOGLA 537
 QY 503 DRYFRALSOETGETIMOTRLATVASGOAISYEVDGMOYVIAAG 546
 Db 538 NGEPHAYDATNGSDLFHFAADSGIAPPVYTLANGKOYVAVEVG 581

RESULT 2
 DHET ACEPO STANDARD; PRT; 738 AA.
 AC P28036;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE ALCOHOL DEHYDROGENASE [ACCEPTOR] PRECURSOR (EC 1.1.99.8).
 GN ADHA.
 OS Acetobacter polyoxogenes.
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Acetobacter.
 OX NCBI_TaxID=439;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=NB11028;
 RX MEDLINE=91159482; PubMed=2001402;
 RA Tamaki T., Fukaya M., Takemura H., Tayama K., Okumura H.,
 RA Kawamura Y., Nishiyama M., Horinouchi S., Bepko T.,
 RT Cloning and sequencing of the gene cluster encoding two subunits of
 RT membrane-bound alcohol dehydrogenase from Acetobacter polyoxogenes.";
 RL Biochim. Biophys. Acta 1088:292-300(1991).
 CC -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL).
 CC -1- CATALYTIC ACTIVITY: PRIMARY ALCOHOL + ACCEPTOR = ALDEHYDE +
 CC REDUCED ACCEPTOR.
 CC -1- COFACTOR: POQ AND HEME.
 CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
 CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
 CC COMPLEX.
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC

CC SPACE (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D00635; BAA00528.1; -
 CC PIR: S14270; S14270.
 CC HSP: P22619; 2MTA.
 CC InterPro: IPR000345; -
 CC InterPro: IPR001479; -
 CC InterPro: IPR002372; -
 CC Pfam: PF01011; Bacterial_POQ_6.
 CC PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 CC PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 CC PROSITE: PS00190; CYTOCHROME_C; 1.
 CC DR PROSITE: PS00364; BACTERIAL_POQ_1; 1.
 CC DR PROSITE: PS00190; CYTOCHROME_C; 1.
 CC KW Oxidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.
 CC FT SIGNAL 1 35
 CC FT CHAIN 36 738 ALCOHOL DEHYDROGENASE [ACCEPTOR].
 CC FT BINDING 650 650 HEME (COVALENT) (BY SIMILARITY).
 CC FT BINDING 653 653 HEME (COVALENT) (BY SIMILARITY).
 CC FT METAL 654 654 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 CC SQ SEQUENCE 738 AA; 80840 MW; 1E2B6ED7BCD92AF6 CRC64;

Query Match 18.3%; Score 564.5; DB 1; Length 738;
 Best Local Similarity 28.0%; Pred. No. 5.7e-32;
 Matches 165; Conservative 85; Mismatches 261; Indels 79; Gaps 14;

QY 10 SAGALALLAARFAQVTVTDE-----LLANPAGEWISYCONENRHSPLTQIT 61
 Db 17 TAGTICALISGYATWASADGQATGEALIHADHPNMWMTYGRYSEGRYSPLDQINR 76
 QY 62 ENVGQLOLVWANGMOPGKQVY-TPLIHGVWYLANPGVIOAIDAKGDLIWEHRRLP 119
 Db 77 SNVGNLKLAWYLDLTNRGQEGTPLYIDGVYATTTNMSMAKAVAAIGKILMSTDPYRG 136
 QY 120 NIATLNSFGEPTRGMAIXGTNYEVSMDNHLVALDTATGQVTPVDKRGQEDVNSN- 175
 Db 137 NIADKGCDDTYNRGAAYWNGKVFYETFGRLIALDAKTGLWVSNTLPPAEELGKGRSY 196
 QY 176 --SSGPIVANGVIAGSTCOYSPGCG--FVSGHDSATGEELMRNYFLPRAGEEDTWCN 226
 Db 197 TYDGAPRIAKRVIIGN--GGSEFGARGVAFPAETGKVDWRFETAPNPKNEPDHNASD 254
 QY 227 -----ETWGNQDYARMMTGA-----WGLTYDPVTNLVHGSTAVGPAPESTGRTGCT 269
 Db 255 SVLMNKAVQWSPGTGAMTRGGGGTVDMSIYDVPADLVYLGVNGSPWNYKYRSEKGD 307
 QY 270 GTPGGLTGTWTRFAVRDPTGEIYWRHQLTPRNDWDECTFEKMAVTNVQSTEMEGLO 329
 Db 308 SEKGDNLFLGSIYALKRTEGEYVWHQETPRMDQFTSVQOITMLDLP----- 357
 QY 330 SINPNAATGERRVLTGVPCKGTGTMQOPDAETGEFLMARDNTYONMIESID-ENGIVTWNEDATLKE 388
 Db 358 -----NGETRHHVIVHARKNGFEYIIDAKTGEFISGKIVYVWMAAGDLPKGRPIYNP 410
 QY 369 DALIKELDEVDVCTFLFGGRDMSAALNPSGSIYFPLNNVCYDMAAVDOEFT-SMDVY 447
 Db 411 DALYTLTGKEMYGIPGDLGGHNPAAAFSPKGTGLYIIPAQVPEFLTNGVGTTPRDPDSMNLGLDM 470
 QY 448 NTS---NVTKLPPG-----KDMIGRIDALISTGRTLMSVERAANYSPVLSTGGCVLENGGT 496
 Db 471 NLGIDMKNVGIIPDPSPKQAFVVDLKGWIVAMPDOKQAEAMRVDRHGRPNWNGILLATGDLFLOGLA 530
 QY 497 LFGSTDRYFRALSOETGETIMOTRLATVASGOAISYEVDGMOYVIAAG 546
 Db 531 LFOGLANGEPHAYDATNGSDLFHFAADSGIAPPVYTLANGKOYVAVEVG 580


```

RESULT 3
DHET_GLUSU STANDARD; PRT; 757 AA.
ID DHET_GLUSU 005542;
AC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ALCOHOL DEHYDROGENASE [ACCEPTOR] PRECURSOR (EC 1.1.99.8) (G3-ADH
  SUBUNIT 1).
GN ADHA.
OS Gluconobacter suboxydans.
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
OC Gluconobacter.
OX NCBI_TaxID=443;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 36-50.
RC STRAIN=IPO.12528;
RX MEDLINE=97208225; PubMed=9055427;
RA Kondo K., Horinouchi S.;
RT "Characterization of the genes encoding the three-component membrane-
  bound alcohol dehydrogenase from Gluconobacter suboxydans and their
  expression in Acetobacter pasteurianus."
RT Appl. Environ. Microbiol. 63:1131-1138(1997).
RL Appl. FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL)
  -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL)
  -1- CATALYTIC ACTIVITY: PRIMARY ALCOHOL + ACCEPTOR = ALDEHYDE +
  -1- COFACTOR: POQ AND HEME.
  -1- SUBUNIT: HETEROTRIMER (DEHYDROGENASE, CYTOCHROME AND PROTEIN
  ADHS). THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX.
  -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
  SPACE (POTENTIAL).
  -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
  -----
  CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
  CC the European Bioinformatics Institute. There are no restrictions on its
  CC use by non-profit institutions as long as its content is in no way
  CC modified and this statement is not removed. Usage by and for commercial
  CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
  CC or send an email to license@isb-sib.ch).
  -----
  CC EMBL; D86375; BAA19753.1; -
  DR HSSP; P38539; 4AAH.
  DR InterPro; IPR000345; -
  DR InterPro; IPR001479; -
  DR InterPro; IPR002372; -
  DR Pfam; PF01011; Bacterial_POQ_6.
  DR PROSITE; PS00363; BACTERIAL_POQ_1; 1.
  DR PROSITE; PS00364; BACTERIAL_POQ_2; 1.
  DR PROSITE; PS00190; CYTOCHROME_C; 1.
  KW Oxidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.
  FT SIGNAL 1 34
  FT CHAIN 35 757 ALCOHOL DEHYDROGENASE [ACCEPTOR].
  FT MOD_RES 35 35 PYRROLIDONE CARBOXYLIC ACID.
  FT AC_SITE 342 342 BASE (POTENTIAL).
  FT BINDING 653 653 HEME (COVALENT) (BY SIMILARITY).
  FT BINDING 656 656 HEME (COVALENT) (BY SIMILARITY).
  FT METAL 657 657 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
  SEQUENCE 757 AA; 82968 MW; 39B9F90E3B947581 CRC64;

```

```

Query Match 17.9%; Score 553.5; DB 1; Length 757;
Best Local Similarity 27.4%; Pred. No. 3.5e-31;
Matches 168; Conservative 97; Mismatches 253; Indels 93; Gaps 19;

```

```

QY 7 IWSAGATALLAA--PAAQOV---TPVTDLLANPAGEWISYGGONQNRHSPLTQITTT 61
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 16 LLSCAALAFSAAYVAFAQDEGTGTAITSNDGHP-CDWLSTYGRSYSEGRSPLDQINT 74
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 62 ENVGQLQLVWANGQPGKVQV-TPLIHGVMTLANPGVDTQADAKTGLDILWEHRQLP- 119
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

```

DB 75 ENVGKILAMHEDLTNRGQESTPLIVGVYATTNWMSKMAALDAATKLLMSYDPKYPG 134
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 120 NATIANSFGEPTRGALXGNTVYFVSMNHLVADDTAGOVTFVDROGQEDVSN----- 175
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 135 NIADRGCCDVTYSGAAYWNGKVFYFSTGRLIALDAKTKGLVMSYTTTPKENQLOHQSY 194
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 176 --SSGPIVANGVIVAGSTCOYSPFGC--FVSGHDSATGEELMRNYFIPRAGEED----- 226
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 195 TYDGNAPRIAKGVLIGN--GGAEEFARGFVSADFETSKLDMRFETVNPENKPGDASD 252
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 227 -----ETWGNDYEARMTG---AMQITDPTVNLVHGSTAVGPASETQRTPGGT 275
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 253 DILMSKAYPTWKNAMKQGGGTFWDSLVDVPTDLVLGVNGSPWNTKFNSEKGD 312
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 276 LYGNTFRFAVRPDGTGEIYVRHOTLPRDMMDOCEPEMMVTNVDPSTMEGLSINPNA 335
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 313 NLPLGSIVAINPDYTKYVHNPQETPMDEMDYTSVOQIMTLDMPV----- 356
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 336 ANGE--RRVLTVGVCPTGTGMQFPAETGEFLMARDTNYONMIESIDE-NGIVTVEDAILK 393
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 357 -NGEMRHIVVAP-KNGFYIIDAKTKFITGKPYTEKWMANGLDPVIGRPVYVDALMT 414
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 394 ELIDVEYDVCPTLGGRDWPSAALNPDGSIYFPLNNV-----CYDMAAVDQE 440
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 415 LYGKPLWGLPGEIGLGHNFAMAYSPKTKLVYIPAOQIPLLYDGOKGGERAYHDANNLGLD 474
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 441 FTSMDVYNTSNVTKLPFGDMIGRIDADISTGRTL-NSVERAANYS-----PVL 490
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 475 MNKIGLFDNDPEHVAKKDF-----LKLKGLWYVAMPPEKMAAFITNHGPNNGGLL 528
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 491 STGGVGLFNGGDRYPRALSOETGETLMOTRLATVYASGAISIEVDGMOYAI----- 543
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 529 ATAGNVTFQGLANGEPHAYDATNGNDLSFPASALIAAPPVYTAGKQYVAEVWGSI 588
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 544 ---AGGVSYSYG 553
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 589 YPRLVGVGAVRTSG 601
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

```

RESULT 4
DHET_ACAC STANDARD; PRT; 742 AA.
ID DHET_ACAC P18278;
AC 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE ALCOHOL DEHYDROGENASE [ACCEPTOR] PRECURSOR (EC 1.1.99.8).
GN ADHA OR ADH1.
OS Acetobacter aceti.
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
OC Acetobacter.
OX NCBI_TaxID=435;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 36-50.
RX MEDLINE=89255070; PubMed=2722742;
RA Inoue T., Sunagawa M., Mori A., Imai C., Fukuda M., Takagi M.,
  Yano K.;
RT "Cloning and sequencing of the gene encoding the 72-kilodalton
  dehydrogenase subunit of alcohol dehydrogenase from Acetobacter
  aceti."
RT J. Bacteriol. 171:3115-3122(1989).
RN [2]
RP 3D-STRUCTURE MODELING.
RX MEDLINE=95289964; PubMed=7772016;
RA Gozler G.E., Gilles I.G., Anthony C.;
RT "The structure of the quinoprotein alcohol dehydrogenase of
  Acetobacter aceti modelled on that of methanol dehydrogenase from
  Methylobacterium extorquens."
RT Biochem. J. 308:375-379(1995).
CC -1- CATALYTIC ACTIVITY: PRIMARY ALCOHOL + ACCEPTOR = ALDEHYDE +
  REDUCED ACCEPTOR.
  -1- COFACTOR: POQ AND HEME.

```

CC -1- SUBUNIT: TETRAMER OF NON IDENTICAL CHAINS (DEHYDROGENASE,
CYTOCHROME, AND TWO SMALLER UNKNOWN SUBUNITS).
CC -1- SUBCELLULAR LOCATION: PERIPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb.sib.ch).
CC
DR EMBL; D90004; BAA14058.1; -
DR PIR; J50326; J50326.
DR HSSP; P38539; 4A4A.
DR InterPro; IPR000345; -
DR InterPro; IPR001479; -
DR InterPro; IPR002372; -
DR Pfam; PF01011; Bacterial_PQQ_6.
DR PROSITE; PS00363; BACTERIAL_PQQ_1; 1.
DR PROSITE; PS00364; BACTERIAL_PQQ_2; 1.
DR PROSITE; PS00190; CYTOCHROME_C; 1.
KW Oxidoreductase; PQQ; Heme; Periplasmic; Membrane; Signal.
FT SIGNAL 1 35
FT CHAIN 1 35
FT ACT_SITE 36 742 BASE (POTENTIAL).
FT BINDING 343 343 HEME (COVALENT) (BY SIMILARITY).
FT BINDING 649 649 HEME (COVALENT) (BY SIMILARITY).
FT METAL 652 652 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT METAL 653 653 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 742 AA; 81521 MW; 9C6C9268DB825A CRC64;

Query Match 17.7%; Score 547.5; DB 1; Length 742;
Best Local Similarity 27.8%; Pred. No. 8.8e-31;
Matches 176; Conservative 97; Mismatches 244; Indels 115; Gaps 24;

QY 11 AGALALLAAPACVPTVDEL--ANPPAGEWISYQONENRHSPLQITTEENYQGLQ 68
DB 23 AALPYAAVAPARADQGTGEALIHADHPENMLSYGRYSSEQLDINSNNVGLDK 82
QY 69 LVMARGQOPKQV-VTPLIHDGWTYLANPGDVLQADAKTGLDIWEHRQLP-NIATLNS 126
DB 83 LLAGYTLDTNNGQKATPLVVDGIMVATNTNSKKEALDAAATKLLMDYDPKVINADKGC 142
QY 127 FGEPTRMALYGTIVYVSWDNHLVALDPTATGOV---TEVDROGGEEMVSNSSGPI- 180
DB 143 CDYVNRAGYVNGVYFETGTDGLRYAADATKTKVAAVNTIPADASLGKORSTYVDGAVR 202
QY 181 VANGVIVASTCOYSPFGC--FVSGHDSATGEELMRNYFLPRAAGEGD----- 226
DB 203 VANGVIVLIGN--GGAERGARGFSAFDEATGKLMRFYVPPNNKNEPDHAASNNILMKA 260
QY 227 -ETMGNDEYKRWMT-----GAKQIYDPTNIVHGSYAVGASSETOKTPEGGLY--- 277
DB 261 YKTMGP--KAMVYKGGGGGTWDSLVYDPSLDLY---LAVG-----NGSPNNYKRSR 309
QY 278 --GTN---TRFAVRPDTEIVRHOITLPRDNNDQCTEFEMAVNTVDOVSTEMELQSI 331
DB 310 GIGSNFLFSIVALKPEGTGYVNHPRATPDMDYTSVQGIIMLDMPK----- 358
QY 332 NPNAATGE--RVLTGVPCKTGTMMQDAETGEFLMARDNTYONMIESID-----EN 381
DB 359 -----GEMHNVIVAP-KNGFEYVLDATKGEFLSGKNVYVQWANGDPLRGRPNYND 411
QY 382 GIYTVNEDALIKELDYEDVCPFLGGRMPSSAALNPDGSIYFIPLNNCY----- 432
DB 412 GIYTLG-----KFWYGI--PQPLAHNFMAMAYSPKTHLYIPAHQIPEFYKNQVGF 463
QY 433 ----DMVAVDQETSMDEVNTSNVTKLTPGKDMIGRIDIDISTGTILMSVERAANSP 488
DB 464 KPHADSNVGLDMTKNGLDPTPE-AFTATIKDLHGMLLAMDVPMKMETYWKIDHKPWNNG 522

QY 489 VLSNGGVLENGGDRFRALSGEETMTQTLATVYASQAISYEVDQMYYA----- 542
DB 523 ILATGGDLFLQGLANGEBFAYDATNGSDLYKFPDAQGIAPMTYVSNQYVAEVEGNG 582
QY 543 ----IAGGVSYSGSL---NSALAGERVDSTA 567
DB 583 GIYPISMGVGVRGSRGWTVMNHSYIAAFSLDCKA 614
RESULT 5
EXAA_PSEAE STANDARD; PRT; 623 AA.
ID EXAA_PSEAE
AC 092417;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE QUINOPROTEIN ETHANOL DEHYDROGENASE PRECURSOR (EC 1.1.99.-) (OEH).
GN EXAA OR PA1982.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 17933;
RX MEDLINE=99041560; PubMed=9826187;
RA Diehl A., Wintzingerode F., Goerlich H.;
RT "Quinoprotein ethanol dehydrogenase of Pseudomonas aeruginosa is a
RT homodimer: sequence of the gene and deduced structural properties of
RT the enzyme.";
RL Eur. J. Biochem. 257:409-419(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wedman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reiter J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
RN [3]
RP SEQUENCE OF 1-52 FROM N.A.
RC STRAIN=ATCC 17933;
RX MEDLINE=99173751; PubMed=10075429;
RA Schober M., Goerlich H.;
RT "Cytochrome c550 is an essential component of the quinoprotein ethanol
RT oxidation system in Pseudomonas aeruginosa: cloning and sequencing of
RT the genes encoding cytochrome c550 and an adjacent acetaldehyde
RT dehydrogenase.";
RL Microbiology 145:471-481(1999).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 35-616.
RC STRAIN=ATCC 17933;
RX MEDLINE=20203376; PubMed=10736230;
RA Ketel T., Diehl A., Knaute T., Stezowski J.J., Hoehne W.,
RA Goerlich H.;
RT "X-ray structure of the quinoprotein ethanol dehydrogenase from
RT Pseudomonas aeruginosa: basis of substrate specificity.";
RL J. Mol. Biol. 297:961-974(2000).
CC -1- FUNCTION: OXIDIZE PRIMARY ALCOHOLS AND ALSO ACTS ON SECONDARY
CC ALCOHOL, BUT NOT HIGHLY ACTIVE ON METHANOL.
CC -1- COFACTOR: PQQ AND CALCIUM.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: PERIPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch)

DR EMBL: AJ009858; CA08896.1; -
DR EMBL: AE004624; AAC05370.1; -
DR EMBL: AF068264; AAC79657.1; -
DR PDB: 1FLG; 30-AUG-00.
DR InterPro: IPR002372; -
DR Pfam: PF01011; Bacterial_P00_1; FALSE_NEG.
DR PROSITE: PS00363; BACTERIAL_P00_1; FALSE_NEG.
DR PROSITE: PS00364; BACTERIAL_P00_2; FALSE_NEG.
KW Oxidoreductase; PQQ; Periplasmic; Signal; Calcium; 3D-structure.
FT SIGNAL 1 34
FT CHAIN 1 35 623 QUINOPROTEIN ETHANOL DEHYDROGENASE.
FT DISULFID 139 140
FT SEQUENCE 623 AA; 68123 MW; 32DDE5DF20B291D6 CRC64;

Query Match 16.3%; Score 505; DB 1; Length 623;
Best Local Similarity 25.7%; Pred. No. 6.5e-28;
Matches 164; Conservative 105; Mismatches 276; Indels 92; Gaps 23;

3 PSLMASAGALAL--LAAPFAQVTPVDEL--PPAGEWISYSGQNDENRHSPLT 57
9 PAGLRLPSLHCLAFVALGASAGALAKDTWEDIANDKTTDVLQYGGTHAQKWSPLK 68
58 QTTENVGQLVWARGM---QPGKVQVTPPLIHGWMLANPGVYQATIAKTGDLWE 113
69 QVADNVKFLTPAWMSYSEFDEKORQ-ESQALVSQGVIVTYSRLEFLDAKTCKRLMT 127
114 HRRQJN-----IATNSGEPFGMALGTWVFVSMNHLVADLTAGVTFP---VD 165
128 YNHRLPDTRPCCDVYN-----RGAITGDKVFEGTIDLASVALANKNGKVMKKRAD 181
166 RGGEEDVANSQPIVANG-----VIVAGSTQYSPGCF--VSGHDSATGELMRNFTI 218
182 HGAGYTM--TGAPTYIKDKGTGKVLIIHSS--GDEFGVGRLEFARDPDTGEIWMRPV 237
219 PR-----AGEBDETWGNDYEA-----RWMGCA---WGQIYDVTNLVNHG 257
238 EGHMRLNGKSTVTGDKVAPSWPDRNSPTGKVESWISGGAPWQASAFDAETWIIVG 297
258 STAVGASETORCTPGG-----TLGTNTREAVRPDTGEIWRHROTLPRDMMDECTPE 311
298 AGNPGWNTWARTAKGNGHIDSLT-TSGQYGVDPSSSEVKNFIOHTPDNDAMSSGNNE 356
312 MMTNVAVOPSTEMEGLOSINPN-----AATGERVLTGVPCKTGTW--QFDAETGEFL 364
357 LVLEFKAKDKGIVKATTAHADRNCFYVDRSNGKLQNAFPPVDNITWASHIDIKTG--- 413
365 WADDTYQNMIESIDENGIVTVEMLKELDEYVYCTFLGGRMPSPALNPPSGIYF 424
414 --RPVEREQRPPLBPBG-----OKHKAWEVSPFLGGKWNWDMAYSDQDGLFY 461
425 IPIANNVCYDMAVDOEFTSADVYNTSNVTKLPKGMKIMRIDALIDSTGRITMSVERAA 484
462 VPAHHMKEDYWTPEVSYTGSALGKGFIRKRYDDHVSGLRAMPVDSKRVYWEHKEHLP 521
485 NYSPLVSTGGVLENGDTRFRRALSQETGETIMQTRLATVASGOAISYEDVMQVAYA 544
522 LMGVILATAGNLVFTGTGQYFKAFDAKSGKELMKFQTSGLVSPITWEDGEGYLVGT 581
545 GGGVSYSGSLNASLAGEVDST---AIGNNAVYFALP 578
582 ---VGYG-GAVPLMGDMADLTRPVAQGGSEWFKLP 614

RESULT 6
DHML-PARDE
ID DHML-PARDE STANDARD; PRT; 631 AA.
AC P1293;

DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE METHANOL DEHYDROGENASE SUBUNIT 1 PRECURSOR (EC 1.1.99.8) (MDH LARGE
DE ALPHA SUBUNIT) (MEDH).
GN MOXF.
OS Paracoccus denitrificans.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Paracoccus.
OX NCBI_TaxID=266;
RN [1]
RP SEQUENCE FROM N.A. AND SEQUENCE OF 33-49.
RX MEDLINE=87307969; PubMed=3114231.
RA Harms N., de Vries G.E., Maurer K., Hoogenboezem J., Stoutamer A.H.;
RT "Isolation and nucleotide sequence of the methanol dehydrogenase
RT structural gene from Paracoccus denitrificans."
RL J. Bacteriol. 169:3969-3975(1987).
CC -1- CATALYTIC ACTIVITY: PRIMARY ALCOHOL + ACCEPTOR - ALDEHYDE +
CC REDUCED ACCEPTOR.
CC -1- COFACTOR: PQQ.
CC -1- SUBCELLULAR LOCATION: PERIPLASMIC.
CC -1- MISCELLANEOUS: MDH IS THE MAJOR PROTEIN IN THE CELL DURING GROWTH
CC ON METHANOL. (IN P. DENITRIFICANS MDH CONSTITUTES UP TO 15% OF THE
CC TOTAL CELL PROTEIN).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch)
CC
CC EMBL: M17339; AAA88366.1; -
CC HSPB: P38539; AAAH
DR InterPro: IPR001479; -
DR InterPro: IPR002372; -
DR Pfam: PF01011; Bacterial_P00_7.
DR PROSITE: PS00363; BACTERIAL_P00_1; 1.
DR PROSITE: PS00364; BACTERIAL_P00_2; 1.
KW Oxidoreductase; PQQ; Signal; Methanol utilization; Periplasmic.
FT SIGNAL 1 32
FT CHAIN 1 33 631 METHANOL DEHYDROGENASE SUBUNIT 1.
FT DISULFID 135 136 BY SIMILARITY.
FT DISULFID 418 447 BY SIMILARITY.
FT ACT SITE 335 335 BASE (POTENTIAL).
SQ SEQUENCE 631 AA; 69799 MW; 0934DC93PFC5730B CRC64;

Query Match 16.0%; Score 495; DB 1; Length 631;
Best Local Similarity 26.0%; Pred. No. 3.3e-27;
Matches 166; Conservative 111; Mismatches 259; Indels 102; Gaps 25;

4 TSLMASGALAL--LAAPFAQVTPVDEL--LANPAGEMISYSGQNDENRHSPLQIT 60
12 SSLAAVAVMGALVITTAATA-----NDQLELAKDPA--NMWVGGRDQNAQYSEMDIN 65
61 TENVGOLVW--ARGMPGKVQVTPPLIHGVMYLANP-GDVIOAIDA-KTGDLWEHR- 115
66 KENVQQLRANFSFGVLHGH-EGTFLVVGDRMFIHFPFTTALDNEGKILMOKRP 124
116 RQLPNATLNSFEGPTRGMALYGTN-----VYFVSWNHLVALDTAGVTFVDYRGOG 169
125 KNPARTVACCDVNRRLAWPAGDDQKPLIFRTQLDGHTVAMDAGETGRWIMENSDI 184
170 EEMVANSQPIVANGVYIAGSTQYSPGCFVSGHDSATGELMRNFTI----- 218
185 KVGSTFLTAIPYIKDLVTVGSSGAEIVRGVYTAIVKSGEMRRARFATGDEDELLAED 244
219 -----PRAGEE--GDETWGNDYEARMGTGA---WGQIYDVTNLVNHGSTAVGASETO 268
245 FNAPRPHYGQKNLGLETWEGD---AMKIGGGTNGWGVAYDEVDLEFYGSGNPAWNETM 301

QY	269	REPSEGLTGYTNRFAVRPDVTGIVNRHDTLPDNDMDODETEFMATYNDVDPSTEMEGL	328
Db	302	R--PGGNKM--TAAIMWRERATTEGAFAAYOKPTDEWD-----YAGVAMMALISQDEK	350
QY	329	QSINPNAATGERVRLVGVDPCKTGTMMQFPAETGEFLMARDTNYQNMMIESIDE--NGIVTV	386
Db	351	Q-----GOMKRLTLTHPDNRNGIYVTLDRNTODLISA-----DKMDDVTNMVKEV	393
QY	387	NEDALLKEIDVEY-----DVCTPLGGNDWPSAALNPDSDGIYFIFLNNVCYIMMAV	437
Db	394	QDGTGTPVDPPEGTRMDHKARDICPSAMGYHNQGDSDYDPEKKVYMLGINHICMDEPF	453
QY	438	DOEFTSMDYNTSNVTKLPKPKDM-----IGRIDAIDISTGRIMSEVESAANYSPVLST	492
Db	454	MLPYRAGCFYGVATILMYGPKATAMRAGAGQKAKDALSGEKKMKMERFSWGSTMAT	513
QY	493	GGGVLENGSTDRYFRALSOETGETIMQTRLATVYASGAISYEVDGMOYAI--AGGVS	549
Db	514	AGGLTFYVTLDFEIKRARDSDTDGLLMMKFLPSGVISHPMTYKHQGYVALIMYGVGWPG	573
QY	550	YG-----SGLNSALGERV--DSTALGNMYVVAL	577
Db	574	VGLVFDLADPTAGLGSVGAFKRLQEPFGGGVWFSL	611
RESULT	7		
QXOF	PARDE		
ID	XOXF_PARDE	STANDARD:	PRT: 600 AA.
AC	P29368:		
DT	01-APR-1993	(Rel. 25, Created)	
DT	01-OCT-1996	(Rel. 34, Last sequence update)	
DE	15-DEC-1998	(Rel. 37, Last annotation update)	
DE	PURATIVE DEHYDROGENASE XOXF PRECURSOR (EC 1.1.99.-).		
GN	XOXF.		
OS	Paracoccus denitrificans.		
OC	Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;		
OC	Paracoccus.		
OX	NCBI_TaxID=266;		
RL	[1]		
RA	SEQUENCE FROM N.A.		
RA	Harms N.;		
RL	Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	SEQUENCE OF 380-600 FROM N.A.		
RC	STRAIN=PD 1235;		
RC	MEDLINE=92041583; PubMed=1657873;		
RC	Medline=92041583; PubMed=1657873;		
RA	Ras J., Reijnders W.N.M., van Spanning R.J.M., Harms N., Oltmann L.F.,		
RA	Stouthamer A.H.;		
RT	"Isolation, sequencing, and mutagenesis of the gene encoding		
RT	cytochrome c551 of Paracoccus denitrificans and characterization of		
RT	the mutant strain."		
RL	J. Bacteriol. 173:6971-6979(1991).		
CC	-I- COFACTOR: POQ (BY SIMILARITY).		
CC	-I- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL: U34346; AAC44555.1; -;		
DR	EMBL: M75583; AAA2574.1; -;		
DR	PIR: A41378; A41378.		
DR	HSSP: P38539; AAH.		
DR	InterPro: IPR002372; -;		
DR	Pfam: PF01011; Bacterial POQ: 7.		
KM	oxoreductase; POQ; Signal.		
FT	SIGNAL	1 21	POTENTIAL.
FT	CHAIN	22 600	PURATIVE DEHYDROGENASE XOXF.

FT	ACT SITE	318	318	BASE (POTENTIAL).
50	SEQUENCE	600 AA;	65159 MW;	DCA996F1DBC5A30CE CRC64;
	Query Match		15.8%;	Score 487.5; DB 1; Length 600;
	Best Local Similarity		27.1%;	Pred. No. 1e-26;
	Matches 173;		Conservative	96; Mismatches 225; Indels 115; Gaps
QY		10	SAGALALLAPAFQVPTVDELANP-----PAGMISYGCNQNRYHSPLOTITTE	62
DB		6	NGACIALIMSGTAA-----LANEORAGDRDRAQAPMAIQMDVYNTFSTLDDIND	56
QY		63	NVGOLOLW--ARGMOPGKVQVPTLHCOSVWLANP--GDVIOAQIDAK--TGDILWEHR--RO	11
DB		57	NVKDLRAAMTFSTGVLRGH--EGSPLYIGDVYVHTPPENRRFALDLNDCKITLMREPOQ	11
QY		118	LPIATLINSFEGPTRGMALYGTNYVFVSDNHLVALDTAGQVTFDV--DRSGEDMVS	17
DB		116	DPNIVAMSCDQVYRGISVADGMILLGGADPTVVALADATSEVWKSRIQDPGIGETLLA	17
QY		175	NSSGPIVANGYIAGSTCOYSPGCFVSGHDA----TGEELMRNFI-----	21
DB		176	TVVP--VKDVLVIGISGEGY-----VGRMTALMLTLDGSEAMKAMSTGDEELLDVDET	22
QY		219	-----FRABEEDDETWNQDIARWMTGA--WQIYIDPTNLVHGTSNAPASERTOR	26
DB		229	TTHLGKPIGADSSILNSWED--QMOIGGGTITWGFSTDPDLNLYVYGTSNPSTMPSSOR	28
QY		270	GTPEGTLYGNTPEFARPRPTGETIWRHQTLPDNNWDOECFEEMVNTN--VDVOPSTEMG	32
DB		286	--PDDNKM--SMTLMADADDTGAKMKFYQMTPEHDEDDVGVNEMILTIQVYDGO-----	33
QY		328	LQSLNPAATGERRVLTVGPCKTGIMQFQDAEDEFIWARDJTYQNMIE----SIDENG	38
DB		336	-----ERKLLTHFD--RNGLAYTMDRETELLVAE--KYDPVAMNTTGVDMDPS	38
QY		383	IVTVNEAAILKELDVEYD-----VCPFLGGRDMPSAALNPDGSIYFIPLANVCIDM	43
DB		382	-ETYGPRVAAVEASTAONEDENTTGWEPALGTRKDQPAFSPKTNLEVPFNHVCMDY	44
QY		435	MAVDOEFTSMDEV--WTSNVTKLPKPGKMICIRIDAIDISTGRILMSVERAANYSPVLST	49
DB		441	EPFVAVATTAQPIYVGATTLSTMYPAFSPHSGMKNFIAHNTTIGELIKSVYPROFWSGALAT	50
DB		501	AGDVVEFGTLEGYLRKRVADQGTGEELYKRTKFTSGIIGVMVYEHGKQYVIGILSGVGMAG	56
QY		548	VSYGSGILNS-----ALAGERVDS--TAIGNAVYVFLP	57
DB		561	IGLAAGLTNPNDGLAGVGYASLSQYTELGQQLTVFELP	599
	RESULT	8		
	DHML_METOR		STANDARD;	PRT; 626 AA.
AC		AC	P15279;	
DT		DT	01-APR-1990 (Rel. 14, Created)	
DT		DT	01-APR-1990 (Rel. 14, Last sequence update)	
DT		DT	15-DEC-1998 (Rel. 37, Last annotation update)	
DE		DE	METHANOL, DEHYDROGENASE SUBUNIT 1 PRECURSOR (EC 1.1.99.8) (MDH LARGE	
GN		GN	ALPHA SUBUNIT) (MEDH).	
GN		GN	MOXF.	
OS		OS	Methylobacterium organophilum XX.	
OC		OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group.	
OC		OC	Methylobacterium group; Methylobacterium.	
OX		OX	NCBI_TaxID=412;	
RN		RN	1]	
RP		RP	SEQUENCE FROM N.A., AND SEQUENCE OF 29-43.	
RC		RC	STRAIN-ATCC 27886;	
RX		RX	MEDLINE=89008094; PubMed=2459109;	
RT		RT	Michlin S.M., Hanson R.S.;	
			"Nucleotide sequence and transcriptional start site of the	

RT Methylobacterium organophilum XX methanol dehydrogenase structural
 RT gene.
 RL J. Bacteriol. 170:4739-4747(1988).
 CC -1- CATALYTIC ACTIVITY: PRIMARY ALCOHOL + ACCEPTOR = ALDEHYDE +
 CC REDUCED ACCEPTOR.
 CC -1- COFACTOR: POO.
 CC -1- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
 CC OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
 CC MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M22629; AAA50289.1; -
 CC HSSP: P38539; 4AAH.
 CC InterPro: IPR001479; -
 CC InterPro: IPR002372; -
 CC Pfam: PF01011; Bacterial_POO_7.
 CC PROSITE: PS00363; BACTERIAL_POO_1; 1.
 CC PROSITE: PS00364; BACTERIAL_POO_2; 1.
 CC KMOxidoreductase; POO; Signal; Methanol utilization; Periplasmic.
 CC FT CHAIN 1 28 METHANOL DEHYDROGENASE SUBUNIT 1.
 CC FT DISULFID 130 131 BY SIMILARITY.
 CC FT DISULFID 413 442 BY SIMILARITY.
 CC FT ACT_SITE 330 330 BASE (POTENTIAL).
 CC FT SEQUENCE 626 AA; 66677 MW; 8766F6B8371E5DF CRC64;
 SO
 Query Match 15.5%; Score 477.5; DB 1; Length 626;
 Best Local Similarity 26.0%; Pred. No. 5.4e-26;
 Matches 163; Conservative 112; Mismatches 267; Indels 85; Gaps 25;

QY 10 SAGALALLA-AP-AFAOVTVPTDELAPPAGE-WISYQNGQENRHSPLTQTTENVQ 66
 DB 7 SVSALAMALAPALASVAAVANDKVELSKSDNMVMPGRNYSNNSELKOVNKSVMQ 66
 QY 67 LQIYW-ARGQOPKQVQVPLIHGVWYL--ANPGVIOAIDA-KTGDILWEHR-ROLPN 120
 DB 67 LRPAMTSTGLLNH-EGAPLVVDGAMVHTSPNNTF-ALDDDPCHILMOKPKONPA 124
 QY 121 IATNSFGEPTRGALYGTN-----YFVSWDNHVALDPTATGQVTFVDRGGEDMVS 174
 DB 125 ARVACODIVARGLAYWPGDKTPALLIKTQDLRHVVALNAETGEIVKVENSDIKVST 184
 QY 175 NSSGPYANGVYVAGSTCQVSFPECFSGHDSANGELMRYFL----- 218
 DB 185 LTIPAYVKDVIIIGSSGALGVRYLTAYDVKTGGVWRAYATGPKDLLADDEVNKN 244
 QY 219 PRAGEE-GDEFTGNDYEARMTGA---WQIYDPTNLVHYGSTVAVPASETORSTPG 273
 DB 245 AHYGKIGCTTWTEGD---AMKIGGINWGYAYDPGTNLIFETGNPAPANNEMR--PG 299
 QY 274 GLYGTNTRFAVRBDTGEIVMRHQTLPEDNDODECTEFEMVTVNDVOSTMEGLQSI 333
 DB 300 DNKM-TMTIFGRDADTGEAKFGYOKTPHDEWDYAGVNM-----PSEQKD----- 344
 QY 334 NAATGERVVLGVCKTGMQFADGEFLMAR---DTNNQNMIESIDENGITYVEDA 390
 DB 345 --KDKTKRLLTHPDNRNIYVTLDTGALVSANKLDTF--VNVFKYVDLTKGPVPADPE 400
 QY 391 ILKELD-VEYDVPFLGGRWPSAALNPDGSIYFIPLANNYCYMMAVDQETSMDYNT 449
 DB 401 YTRNDHLAKDYCPAMKYHNQGHSDYDPKRELFFMGINHCMDWEPEMLYRAGGFVVG 460
 QY 450 SNVTKLPCK-----DMIGRIDALIDISTGRTLSVERAANYSFVLSTGGGVLFNGSTD 503

DB 461 ATLNNYPCPKGRQNYEGAIKAYNAITGSYKWEKMERFAVWGGTLATADLIFYETLD 520
 QY 504 RYFRALSETGETIMQFLATVASGOAISYEVDGMOYAI---AGGVSYS----- 551
 DB 521 GYLKARSDTDGLMKFKFIPSGAIGYPMPTTHKGTQVAYIYGVGMPGVGLVVDLADPT 580
 QY 552 SGUNSLAGERV-DSTAIGNAVYFAL 577
 DB 581 AGLGAVGAFKRIANTYQGGGVIVFSL 607
 RESULT 9
 ID DM1 METEX STANDARD; PRT; 626 AA.
 AC P16027;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE METHANOL DEHYDROGENASE SUBUNIT 1 PRECURSOR (EC 1.1.99.8) (MDH LARGE
 DE ALPHA SUBUNIT) (MDH).
 GN MOXF.
 OS Methylobacterium extorquens.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Methylobacterium group; Methylobacterium.
 OX NCBI_TaxID=408;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AM1 / NCIB 9133;
 RX MEDLINE=90337342; PubMed=2116368;
 RA Anderson D.J., Morris C.J., Nunn D.N., Anthony C., Lidstrom M.E.;
 RT "Nucleotide sequence of the Methylobacterium extorquens AM1 moxf and
 RT mox genes involved in methanol oxidation.";
 RL Gene 90:173-176(1990).
 RN [2]
 RP SEQUENCE OF 28-53.
 RC STRAIN=AM1 / NCIB 9133;
 RX MEDLINE=89350892; PubMed=2504152;
 RA Nunn D.N., Day D., Anthony C.;
 RT "The second subunit of methanol dehydrogenase of Methylobacterium
 RT extorquens AM1.";
 RL Biochem. J. 260:857-862(1989).
 RN [3]
 RP DISULFIDE BONDS.
 RX MEDLINE=95384759; PubMed=7656012;
 RA Blake C.C.F., Ghosh M., Harlos K., Avezoux A., Anthony C.;
 RT "The active site of methanol dehydrogenase contains a disulphide
 RT bridge between adjacent cysteine residues.";
 RL Nat. Struct. Biol. 1:102-105(1994).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (1.94 ANGSTROMS).
 RX MEDLINE=95253818; PubMed=7735834;
 RA Ghosh M., Anthony C., Harlos K., Goodwin M.G., Blake C.;
 RT "The refined structure of the quinoprotein methanol dehydrogenase
 RT from Methylobacterium extorquens at 1.94 A.";
 RL Structure 3:177-187(1995).
 CC -1- CATALYTIC ACTIVITY: PRIMARY ALCOHOL + ACCEPTOR = ALDEHYDE +
 CC REDUCED ACCEPTOR.
 CC -1- COFACTOR: POO.
 CC -1- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
 CC OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
 CC MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

DR EMBL; M31108; AAA25380.1; .
 DR PIR; J07908; S07908.
 DR PIR; J07906; J07906.
 DR HSSP; P38539; 4AAH.
 DR InterPro; IPR001479; .
 DR InterPro; IPR002372; .
 DR Pfam; PF01011; Bacterial_PQQ_7.
 DR PROSITE; PS00363; BACTERIAL_PQQ_1; 1.
 DR PROSITE; PS00364; BACTERIAL_PQQ_2; 1.
 DR Oxioreductase; PQQ; Signal; Methanol utilization; Periplasmic.
 FT SIGNAL 1 27
 FT CHAIN 28 626 METHANOL DEHYDROGENASE SUBUNIT 1.
 FT DISULFID 130 131
 FT DISULFID 413 442
 FT ACT_SITE 330 330
 SQ SEQUENCE 626 AA; 68434 MW; 64988DDAFDAD34C CRC64;

Query Match 15.2%; Score 468.5; DB 1; Length 626;
 Best Local Similarity 25.4%; Pred. No. 2.3e-25;
 Matches 159; Conservative 113; Mismatches 271; Indels 83; Gaps 23;

QY 10 SAGALALLA-AP-AFAQVYVTDELLANPAGE-WISYQNGENRHSPLTQITTEVGO 66
 DB 7 SVSALMLALAPALSGAYANDKVELSKSDNWMVMPKNSNNSDLKQINKGNVAK 66
 QY 67 LQIYV--ARGMOGKVVPFLIDGVYKL--ANDGVIAIDAKTGDLLMEHR-RQLPNI 121
 DB 67 LRPAMTFTSLILNGH-EGADLVVDGKMTHTSPNPFALGLDDPTILMQDPKONPAA 125
 QY 122 ATINSEPTFGMALYGTN-----YFVSWDNHLVALDTATQVTFEDVDRGGEDEVSN 175
 DB 126 RAVACCDLVNRGLAYMGDGTALILKTLQDGNVALNAEIGEYWKVKNENSIKRGSL 185
 QY 176 SSGPIYANGVIYAGSTQYSPFCFVSGHDSATGEELMNFYI-----P 219
 DB 186 TAPVYVKKDVIILSSGAEELGVYLTAVDKTGEQVWRAVATGPRKDLLASDFENIKRP 245
 QY 220 RAGEE--GDETMGNDYEARMVTA--WGQITDPTNLVHGSTAVGPAFSGTGTGG 274
 DB 246 HYQKGGTGTGEGD--AMKIGGTTNGVYADPCTNLITFTGTPAPWNETMR--PGD 300
 QY 275 TLYGTTREAVRPDGEIYWRHOTLPRDWMDOCTEEMVNTVNDVQSTEMEGLSINPN 334
 DB 301 NKW-TWTIGRADTGEAFGYOKTPHEDM-----YAGVNVMLSEOKD----- 344
 QY 335 AATGERRVLTGVPCCTGMQPDATGEELMAR--DTYQNMIESIDENGIVTANEDAI 391
 DB 345 -KGGKARKLLTHPDRNGIYITLDRDGLVLSAKLLDT--VNVFSSVDLKTGPVARDPEY 401
 QY 392 LKELD-VEYDVCPFLGGRDMPSSALNPDGIVFTPLNNVCYDMMAVADQETSMVYNTS 450
 DB 402 GTFMDHLADICPSANGYHNGHSDYDPRKLEFIMGINICDMWEPMLPYRAGQEFVGA 461
 QY 451 NWTKLPPGK-----DMIGRIDAIIDISTGRTLSVERAANYSPVLSTGGVLENGGTD 504
 DB 462 TTNMVRPGPGDRQVNEGLOIKAYNAITGDYKWEKMERFAVGMATAGDILVFGYGLDG 521
 QY 505 YFRALSOEGETILMOTRLATVASGOAISYEVDMQVYAL--AGGVSYS-----S 552
 DB 522 YLKARSDTGLDLIMKFKIPSGAIGYPMTHKGTQVAILYVGMPGVGLVADLADPTA 581
 QY 553 GUNSLAGERV-DSTAIGNAVYVFAI 577
 DB 582 GLGAVGAFKKLANYTQMGGVVVFSL 607

RESULT 10
 ID DHML_METME STANDARD; PRT; 571 AA.
 AC P38539;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE METHANOL DEHYDROGENASE SUBUNIT 1 (EC 1.1.99.8) (MDH LARGE ALPHA
 DE SUBUNIT) (MEDH).
 OS Methylobacillus methylotrophicus (Bacterium W3A1).
 OC Bacteria; Proteobacteria; beta subdivision; Methylobacillus group;
 OC Methylobacillus.
 OX NCBI_TaxID=17;
 RP [1]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RX MEDLINE=94059969; PubMed=8241148;
 RA White S., Boyd G., Mathews F.S., Xia Z.-X., Dai W.-W., Zhang Y.-F.,
 RA Davidson V.L.;
 RT "The active site structure of the calcium-containing quinoprotein
 RT methanol dehydrogenase.";
 RT Biochemistry 32:12955-12958(1993).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
 RX MEDLINE=93054513; PubMed=1331050;
 RA Xia Z.-X., Dai W.-W., Xiong J.-P., Hao Z.-P., Davidson V.L., White S.,
 RA Mathews F.S.;
 RT "The three-dimensional structures of methanol dehydrogenase from two
 RT methylotrophic bacteria at 2.6-A resolution.";
 RT J. Biol. Chem. 267:22289-22297(1992).
 CC -1- CATALYTIC ACTIVITY: PRIMARY ALCOHOL + ACCEPTOR = ALDEHYDE +
 CC REDUCED ACCEPTOR.
 CC -1- COPOLYMER: TWO MOLECULES OF PQQ AND TWO MOLECULES OF CALCIUM
 CC PER TETRAMER.
 CC -1- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
 CC OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
 CC MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.
 DR PDB; 4AAH; 08-DEC-96.
 DR InterPro; IPR001479; .
 DR InterPro; IPR002372; .
 DR Pfam; PF01011; Bacterial_PQQ_7.
 DR PROSITE; PS00363; BACTERIAL_PQQ_1; FALSE_NEG.
 DR PROSITE; PS00364; BACTERIAL_PQQ_2; 1.
 DR Oxioreductase; PQQ; Methanol utilization; Periplasmic; 3D-structure;
 KM Calcium.
 FT DISULFID 103 104
 FT DISULFID 379 408
 FT ACT_SITE 297 297
 FT HELIX 2
 FT TURN 11 12
 FT STRAND 14 14
 FT TURN 17 18
 FT TURN 21 22
 FT STRAND 26 27
 FT TURN 34 36
 FT TURN 37 39
 FT STRAND 41 47
 FT TURN 59 61
 FT TURN 62 63
 FT STRAND 64 68
 FT TURN 71 73
 FT STRAND 75 79
 FT TURN 80 81
 FT TURN 83 84
 FT TURN 86 90
 FT STRAND 96 101
 FT HELIX 103 104
 FT TURN 103 104
 FT STRAND 112 114
 FT TURN 115 116
 FT STRAND 117 121
 FT TURN 123 124
 FT STRAND 126 131
 FT TURN 132 134
 FT STRAND 137 142
 FT TURN 146 148
 FT HELIX 149 149
 FT TURN 151 151
 FT STRAND 157 159

BASE (POTENTIAL).

FT	TURND	160	161
FT	STRAND	162	166
FT	STRAND	168	168
FT	HELIX	170	172
FT	TURND	173	173
FT	STRAND	177	182
FT	TURND	183	185
FT	STRAND	188	193
FT	TURND	198	202
FT	STRAND	203	203
FT	TURND	205	210
FT	TURND	212	213
FT	HELIX	219	222
FT	TURND	223	223
FT	HELIX	226	231
FT	STRAND	241	243
FT	TURND	244	247
FT	STRAND	248	252
FT	HELIX	261	263
FT	TURND	269	270
FT	STRAND	273	278
FT	TURND	279	281
FT	STRAND	284	289
FT	STRAND	305	311
FT	TURND	312	313
FT	STRAND	314	322
FT	TURND	324	325
FT	STRAND	327	332
FT	TURND	333	335
FT	STRAND	338	343
FT	TURND	346	347
FT	STRAND	351	355
FT	TURND	356	359
FT	STRAND	360	363
FT	HELIX	365	367
FT	STRAND	370	370
FT	TURND	371	372
FT	STRAND	373	373
FT	STRAND	375	378
FT	TURND	382	383
FT	STRAND	390	390
FT	STRAND	392	394
FT	TURND	395	398
FT	STRAND	399	405
FT	STRAND	407	413
FT	TURND	420	421
FT	STRAND	427	433
FT	TURND	435	436
FT	STRAND	438	439
FT	STRAND	444	450
FT	TURND	451	454
FT	STRAND	455	462
FT	STRAND	469	469
FT	STRAND	471	473
FT	TURND	474	476
FT	STRAND	477	481
FT	TURND	483	484
FT	STRAND	486	491
FT	TURND	492	495
FT	STRAND	496	502
FT	STRAND	512	516
FT	TURND	517	518
FT	STRAND	519	526
FT	TURND	530	533
FT	HELIX	534	538
FT	TURND	539	539
FT	TURND	543	545
FT	HELIX	546	552
FT	TURND	553	555
FT	HELIX	556	558
FT	TURND	559	559
FT	STRAND	565	570
SO	SEQUENCE	571 AA:	62449 MW: 0BE94EA5A2DAB1E1 CRC64

[illegible]

RT "Characterization of the gcd gene from *Escherichia coli* K-12 W3110
 RT and regulation of its expression.";
 RL J. Bacteriol. 175:568-571(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / W3110;
 RX MEDLINE=94261430; PubMed=8202364;
 RA Fujita N., Mori H., Yura T., Ishihama A.;
 RT "Systematic sequencing of the *Escherichia coli* genome: analysis of
 RT the 2.4-4.1 min (110,917-193,643 bp) region.";
 RL Nucleic Acids Res. 22:1637-1639(1994).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474(1997).
 RN [5]
 RP TOPOLOGY
 RA MEDLINE=93286127; PubMed=8509415;
 RX Yamada M., Sumi K., Matsushita K., Adachi O., Yamada Y.;
 RT "Topological analysis of quinoprotein glucose dehydrogenase in
 RT *Escherichia coli* and its ubiquinone-binding site.";
 RL J. Biol. Chem. 268:12812-12817(1993).
 RN [6]
 RP 3D-STRUCTURE MODELING.
 RA MEDLINE=96128046; PubMed=8554505;
 RX Cozier G.E., Anthony C.;
 RT "Structure of the quinoprotein glucose dehydrogenase of *Escherichia*
 RT *coli* modelled on that of methanol dehydrogenase from *Methylobacterium*
 RT *extorquens*.";
 RL Biochem J. 312:679-685(1995).
 CC -1- FUNCTION: GDH IS PROBABLY INVOLVED IN ENERGY CONSERVATION RATHER
 CC THAN IN SUGAR METABOLISM.
 CC -1- CATALYTIC ACTIVITY: D-GLUCOSE + ACCEPTOR = D-GLUCONO-1,5-LACTONE +
 CC + REDUCED ACCEPTOR.
 CC -1- COFACTOR: POO.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X51323; CAA35706.1; -
 DR EMBL: D12651; BAA02174.1; -
 DR EMBL: D26562; CAB20298.1; -
 DR EMBL: AE000122; AAC73235.1; -
 DR PIR: JVO107; JVO107.
 DR HSSP: P38539; 4AAH.
 DR ECGene: EG10369; gcd.
 DR InterPro: IPR001479; -
 DR InterPro: IPR002372; -
 DR Pfam: PF01011; Bacterial_POO_7; 7.
 DR PROSITE: PS00363; BACTERIAL_POO_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POO_2; 1.
 KW Oxidoreductase; POO; Transmembrane; Cytoplasmic.
 FT DOMAIN 1 10
 FT TRANSMEM 11 37
 FT DOMAIN 38 40
 FT TRANSMEM 41 58
 FT DOMAIN 59 62
 FT TRANSMEM 63 81
 FT DOMAIN 82 95
 FT PERIPLASMIC (PROBABLE).

FT TRANSMEM 96 110 PROBABLE.
 FT DOMAIN 111 118 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 119 141 PROBABLE.
 FT DOMAIN 142 146 PERIPLASMIC (PROBABLE).
 FT ACT SITE 466 466 BASE (PROBABLE).
 FT CONFLICT 59 59 R -> L (IN REF. 1).
 FT CONFLICT 149 156 TUSADATP -> HLKRRCHT (IN REF. 1 AND 2).
 FT CONFLICT 193 193 N -> K (IN REF. 1 AND 2).
 FT CONFLICT 666 666 Q -> H (IN REF. 2).
 SQ SEQUENCE 796 AA; 86747 MW; D9EDC05A12894E9 CRC64;

Query Match 12.7%; Score 392.5; DB 1; Length 796;
 Best Local Similarity 24.3%; Pred. No. 6.5e-20;
 Matches 179; Conservative 85; Mismatches 231; Indels 241; Gaps 36;

QY 6 LMAS-----AGALALAPAFNOVPTPELLANPPAGEMISGNOENYRHSPLTQ 58
 DB 135 LTMAGFNDOEINGTISADATPAEA-ISPVDQ-----DPAVYGRNDSQRFSPILKQ 185
 QY 59 ITTENYQLOLVWA-----RGMPGKY--QVPLIHGVYLANPGDVIQADAKTGD 109
 DB 186 INADNVHNLKEAMVFTGDKVQPNPDGETTNEVPIKVGDTLYLCTAHQRLFPALDAAGK 245
 QY 110 LIMEHRRQPLNATLNSPEPT-RGALXGTNVFVS-----WDNHLVAL 153
 DB 246 EKHWDH--PELKTNSFQVHTCRGSYHEAKAETASPEWADCPRRITLPPVNDRLAI 302
 QY 154 DPTAGOV--TF-----DVRQGEDMYNSSGPIVANGVIT-AGS--TCQYS- 195
 DB 303 NANGKLCETFAKNGVLANQSMPTKPG---LKEPISPIITDKITVMAAGSYVDNST 358
 QY 196 -FPGCVSHSDATGEELRNFTFRA-----GEEDDEFWANDYKARMGTGAGQITTY 247
 DB 359 RETSGVIRGEFDVNTGELLMA--FDGAKDPNAPISDEHTFFNS-----PNSAPAPAY 409
 QY 248 DPTNVLVHYGTAAGPASPSTGRTGPGTLYGNTREFAVRPPGTGELVWRHQTLPDNDQD 307
 DB 410 DAKLDLVYLPKMTPTTPTDIMGKNTREBOERYAST-LALNATTKLANSYQVHHDLMDMD 468
 QY 308 CTFEMVATNVQVSTEMEGLOSINPNAATGERVLTGVPCKTGMQFQAEETGFL-- 364
 DB 469 -----LPAQPT-----LADITVN--GQKVVIYAPAKTGNIFFLDRNGELVVA 511
 QY 365 -----WARDNYQNMIESIDENCI-----VYNEDAILKEL--DVEDD-- 400
 DB 512 PEKVPQGAAGADVTTPQPESELSFRPTKDLGADMGATMFDQLCVRFHQMREGI 571
 QY 401 -----VCPFLGGRDWPSSALNP----- 418
 DB 572 FTPESEQGTIVFPGSLGMEFWGGISVDPNREVAITANPAPLFFVSKLIPRGGNPMQPKD 631
 QY 419 -----DSGI-----YFIPLVNVCYDMAVDOEFTSMQVNTSNVTKLPPGKDMIGRI 465
 DB 632 AKGTGTESGIQPOYGVPPYVTLN-----PFLS-----PFGJLPCPKPANGYI 672
 QY 466 DAIDISTGRTLSYVERAANYS-----PY-----ISTGGVLFNCGT--DRY 505
 DB 673 SALDKTNEVYWKKRIGTPQDSMPFPMVVPVFFPMKGMPLGPISTAGVLETAATADNTY 722
 QY 506 FRALSQETGETLMQTRLATVASGA--ISYEDGMQYVAI-AGGVSYGSLNSALAGER 562
 DB 723 LRAYMNSGEKLMQGRLP--AGGATPMTEVNVNGKVYVVISAGHGSFG----- 779
 QY 563 VDSRAIGNAVVEPALP 578
 DB 780 ---TRMGDIYVVALP 792
 RESULT 12
 ID DHGA.ACICA STANDARD; PRT; 801 AA.
 AC P05465;

DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE GLUCOSE DEHYDROGENASE-A [PYRROLOQUINOLINE-QUINONE] PRECURSOR
 DE (EC 1.1.99.17) (QUINOPROTEIN GLUCOSE DH) (GDH-A).
 GN GDH.
 OS Acinetobacter calcoaceticus.
 OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 OC Acetobacter.
 OC NCB1_TaxID=471;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LM2 79.41;
 RX MEDLINE=86289368; PubMed=3399393;
 RA Cleton-Jansen A.-M., Goosen N., Odle G., van de Putte P.;
 RT "Nucleotide sequence of the gene coding for quinoprotein glucose
 dehydrogenase from Acinetobacter calcoaceticus."
 RL Nucleic Acids Res. 16:6228-6228(1988).
 CC -1- CATALYTIC ACTIVITY: D-GLUCOSE + ACCEPTOR = D-GLUCONO-1,5-LACTONE +
 CC + REDUCED ACCEPTOR.
 CC -1- COFACTOR: POO.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC.
 CC -1- MISCELLANEOUS: ACINETOBACTER CALCOACETICUS CONTAINS TWO DIFFERENT
 CC POO DEPENDENT ENZYMES WITH GDH ACTIVITY. GDH-A PREFERENCES
 CC 2-DEOXYGLUCOSE AS SUBSTRATE, THE SPECIFIC SUBSTRATES FOR GDH-B ARE
 CC DISACCHARIDES.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X07235; CAA30222.1; -;
 DR PIR: S00943; S00943.
 DR InterPro: IPR001479; -;
 DR InterPro: IPR002372; -;
 DR Pfam: PF01011; Bacterial_POO_7.
 DR PROSITE: PS00363; BACTERIAL_POO_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POO_2; 1.
 KM Oxidoreductase; POO; Transmembrane; Periplasmic; Signal.
 FT SIGNAL 1 801
 FT CHAIN ? 801
 FT TRANSMEM 39 55
 FT TRANSMEM 59 79
 FT TRANSMEM 94 108
 FT TRANSMEM 119 138
 FT ACT SITE 471 471
 FT ACT SITE 471 471
 SEQUENCE 801 AA; 86956 MM; 2F4E222E35450D8 CRC64;
 Query Match 12.6%; Score 390; DB 1; Length 801;
 Best Local Similarity 23.4%; Pred. No. 9.8e-20;
 Matches 169; Conservative 88; Mismatches 226; Indels 240; Gaps 33;
 QY 20 PAFQAVPYVDELANPPAGMISYIGONQENYRSPITOTTENVGLOGLVMARQMPGK 79
 DB 152 PETAQAVGVAAE-----SDMPAYGRTQAGVRSPLKQINDQVWKDLKVAWT--LRTGD 202
 QY 80 V-----QVTPLIHGVMIANPGBVIOAIDAKTGDLMEHRQLPNTATLMSFG 128
 DB 203 LKTNDSEGTNTQYTPKIGNNMFTCTAHOQLAIDATGKRR--RDPKLTGDSFG 239
 QY 129 EPT-RGMALYGTN-----VYFVSMNDHLVALDTATGQVTFD 165
 DB 260 HLTGRGVYTDANNTTEFANSLSQKSSSTQCPKRVFVPPVNDGRVLAVNADTGACITDFG 319
 QY 166 RGQGDWMS-----NSSGPVIANGVIVAGS--TCQYS---PFGGVSHDSATG 209

DB 320 QMGVNIQEFMPYAPYPGYNPTSGIVTGSTVIVAGSVTNNYNSKBPSC-VIRGYDVNTG 378
 QY 210 EELMRNFETPRAGE-----EGDETGNDEYARMNTGAMQIYDDPTNLVHGSRAVGP 263
 DB 379 KLMV-VFDIADPNAMPBEGTTFVHNS-----PNAMPVLIADKLDIV-VYPTGV-- 427
 QY 264 ASETORCTPGGTLYG-----TNTFPAVRDGEIYWRHQTLPRNDMDCETFEEM 313
 DB 428 -----GTP--DIWGDRTTELKERYANSMLAINASTGKLWNQOTHTHDLMDMDVPSQPS 479
 QY 314 VTNVDPVOSTEMEGLOSINP--NAATGERVLTGP-----CKTG--TMQFDP 357
 DB 480 LADIKNKAGQTVPAIVLTGTGNAFVLDRR--NGQPIVPTERPVQTVARQGTGEFY 537
 QY 358 AENGEE-----LMA-----ROTNYNMLESIDENIYVNE 389
 DB 538 SKTQPFSDLLAPQDKLTQDKMGMATMLDQMCVRSFKRLNYGIYTPSENGTL----- 592
 QY 390 ALIKELDEVYDVCPTFLGGRDMSAALNPS-----GIYFI----- 425
 DB 593 -----VFGNIGVEFMGMSVNPDRQYAVANPISGLPFSRLIPADPNRAQTAG 641
 QY 426 -----PLNNVCYDMAVDOEFTSMDYNTSNVTKLPPEKDMIGRIDAIDISTRTLM 477
 DB 642 AGTEQGVQPMYGVY-----GVEISAFSLPGLCKQAPMGVYAGVDLKTHEVVM 691
 QY 478 -----SYERRAANTSPVLSITGGVLFNGGT-DYFPAALSOETETLM 518
 DB 692 KKRIGTRDSLPNIFQPLPAKIGVPGIGSSIYAGVMPFGAODVRLAFENVTNKKML 751
 QY 519 QTRLATVAGSCA--ISYEDVMQYVAI-AGGGVSYSGSLNSALAGRVSTAGNAVYVF 575
 DB 752 EARLP--AGQQAIPMYEINGKQYVIMAGGHSFG-----TKMGDLVAV 795
 QY 576 ALP 578
 DB 796 ALP 798
 RESULT 13
 DHG_GLUOX STANDARD; PRT; 808 AA.
 ID DHG_GLUOX P271175;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE GLUCOSE DEHYDROGENASE [PYRROLOQUINOLINE-QUINONE] PRECURSOR
 DE (EC 1.1.99.17).
 GN GDH.
 OS Gluconobacter oxydans.
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Gluconobacter.
 OC NCB1_TaxID=442;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92017653; PubMed=1833618;
 RA Cleton-Jansen A.-M., Dekker S., van de Putte P., Goosen N.;
 RT "A single amino acid substitution changes the substrate specificity
 of quinoprotein glucose dehydrogenase in Gluconobacter oxydans."
 RL Mol. Gen. genet. 229:206-212(1991).
 RN [2]
 RP REVISION TO 213.
 RA Goosen N.;
 RL Submitted (FEB-1992) to the EMBL/Genbank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-GLUCOSE + ACCEPTOR = D-GLUCONO-1,5-LACTONE +
 CC + REDUCED ACCEPTOR.
 CC -1- COFACTOR: POO.
 CC -1- SUBCELLULAR LOCATION: OUTSIDE OF THE PERIPLASMIC MEMBRANE.
 CC -1- MISCELLANEOUS: THE P1 FORM CAN OXIDIZE ONLY D-GLUCOSE, WHILE THE
 CC P2 FORM CAN ALSO OXIDIZE MALTOSE. THE SEQUENCE OF P1 FORM IS SHOWN
 CC HERE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: X62710; CAA44594.1; ALT_SEQ.
 DR PIR: S17716; QPKEX.
 DR InterPro: IPR002372; -
 DR InterPro: IPR002372; -
 DR Pfam: PF01011; Bacterial_PQO; 7.
 DR PROSITE: PS00363; BACTERIAL_PQO_1; 1.
 DR PROSITE: PS00364; BACTERIAL_PQO_2; 1.
 KW Oxidoreductase; PQO; Transmembrane; Periplasmic; Signal.
 FT SIGNAL 1 219
 FT CHAIN 720 808
 FT TRANSMEM 35 54
 FT TRANSMEM 59 76
 FT TRANSMEM 94 108
 FT TRANSMEM 123 138
 FT ACT SITE 470 470
 FT VARIANT 788 788 H -> N (IN P2 FORM).
 SQ SEQUENCE 808 AA; 87567 MW; 0F4160DA78652445 CRC64;

Query Match 11.0%; Score 340; DB 1; Length 808;
 Best Local Similarity 21.8%; Pred. No. 3.1e-16;
 Matches 159; Conservative 98; Mismatches 259; Indels 212; Gaps 30;

QY 13 ALALLAFAQVTPYTDLLA-----NPPAGMISYGGNGEYRHSPLTQIT 61
 DB 128 AVLAFAFASLFTDPHDISGLPQIINASPADPDNVPASMAHAYGRQAGDRMSPLNQINA 187
 QY 62 ENVGOLQVW-----ARGMQPKV--OVPLHDGVVYLANPDVQAIQADKGLDI 112
 DB 188 TIVSNLAKVAHHTKDMANSDNPGEDTNTATIEFNNTLYMCSLQKLEAVDGAIGANKW 247
 QY 113 EHRROD--PNTATLN---SFGC--PTRGALYGT-----NVEYVSDNHLVALDT 155
 DB 248 VYDPKLIQINPGRHOLTCRGVSEFHERPANAMSDGNAPFDCAKDILPVN--DGRLEVEDA 306
 QY 156 AIGQVTFDY--DNGGGEIDMSN-----SSGRPIVANGIYA-----GSCQYSP 196
 DB 307 DNGKTCGFGNGGEIDLRVNPQYPTTGGYEPTSPVITDKLIANSATIDNGSVKQASG 366
 QY 197 FCGFVSGHDSATGEELM---RNYFIPRAGEGDETWGNDYEARMMWGAMQITVDPVTN 252
 DB 367 ---ATQAFVYTGKRWYVDASNPNDLPDSHVFHNHNSNSMIVS-----STDANLN 418
 QY 419 LVIYPMGV-----GTPDQMGSDRTKDSERFAAGIALNALADKLFAMFYQTVHNDLM 469
 QY 305 DOBCFPEMANVTNDVOPSTMEGLDSINPNATGERRVLTGPRCTGTMMQDADTEG--EF 363
 DB 470 DMELPSQPSLVNVDOKDGLVPAIYA-----PTVTGDLFVLDRTGKEI 513
 QY 364 LMARDT-----NYQMIESIDENGIVYNEDAILKELDV----- 397
 DB 514 VARPELPVQGAAPRGHTSPQPMQO---LTLRPKNPLNDSDINGTIFDQMFCSIYHT 570
 QY 398 -EYD-----VCPFLGGRDWPASALNPSGIYFPLNVCYDMAVDO----- 439
 DB 571 LREGEFTPLSLKSLIFPGDGMFEMGLAVDPQOVAFANPISLPFVSQLVPRGPNP 630
 QY 440 -----EFTSMDEVYNTSNVTKLPPGKDMT-----GALDAIDIST 472
 DB 631 LWPENAKGTGEGLOHNYGIPYAVNLHPFLDPLVLLPFGIMPKCRTPPMGWVAGIDDKT 690

QY 473 GRTLMSVERAANS-----PV-----LSTGGVLP--NGTDRYFRALS 510
 DB 691 NKVYQWRNGLTRDMSGSSLPPIPIKIGVPSLGGPSTANLGLFRAHMDYIRAVN 750
 QY 511 QETGETLMQTRLATVASGAQISYVDQMVAYAGGVSYGSSLSALNGERYDSTAIGN 570
 DB 751 LTTGKVLMDORLPAGAQATPTFYAINQKQYI-----VTVAGGNSF-----PTRMDD 797
 QY 571 AVYVFAIP 578
 DB 798 DIALYALP 805

RESULT 14
 QUITACIA
 ID QUITACIA STANDARD; PRT; 809 AA.
 AC Q59086;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE QUINATE/SHIKIMATE DEHYDROGENASE [PYRROLOQUINOLINE-QUINONE]
 DE (EC 1.1.99.25) (MAD(P)-INDEPENDENT QUINATE DEHYDROGENASE).
 GN QUITA.
 OS Acinetobacter calcoaceticus.
 OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 OC Acinetobacter.
 OX NCBI_Taxid:471;
 RN [1]
 RP SEQUENCE FROM N.A. AND CHARACTERIZATION.
 RC STRAIN-BD143 / ADPL;
 RA MEDLINE=95095936; PubMed=8002591;
 RA Elismore D.A., Ornston L.N.;
 RT "The pca-pob supraoperonic cluster of Acinetobacter calcoaceticus
 RT contains quia, the structural gene for quinate-shikimate
 RT dehydrogenase.";
 RL J. Bacteriol. 176:7659-7666(1994).
 RN [2]
 RP SEQUENCE OF 1-20 FROM N.A.
 RC STRAIN-BD143 / ADPL;
 RA MEDLINE=96011389; PubMed=7592351;
 RA Elismore D.A., Ornston L.N.;
 RT "Unusual ancestry of dehydratases associated with quinate catabolism
 RT in Acinetobacter calcoaceticus.";
 RL J. Bacteriol. 177:5971-5978(1995).
 CC -1- FUNCTION: CAN ACT EITHER ON QUINATE OR ON SHIKIMATE.
 CC -1- CATALYTIC ACTIVITY: QUINATE + PYRROLOQUINOLINE-QUINONE = 5-
 CC DEHYDROQUINATE + REDUCED PYRROLOQUINOLINE-QUINONE.
 CC -1- CATALYTIC ACTIVITY: SHIKIMATE + PYRROLOQUINOLINE-QUINONE = 3-
 CC DEHYDROSHIKIMATE + REDUCED PYRROLOQUINOLINE-QUINONE.
 CC -1- COFACTOR: PQO.
 CC -1- PATHWAY: QUINIC ACID CATABOLIC PATHWAY; FIRST STEP. THIS PATHWAY
 CC ALLOWS GROWTH OF BACTERIA WITH QUINATE BY ITS CONVERSION TO
 CC PROTOCATECHUATE AND SUBSEQUENT METABOLISM BY THE BETA-KETOADIPIRATE
 CC PATHWAY.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- INDUCTION: BY PROTOCATECHUATE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQO DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: L05770; AAC37161.1; -
 DR Pfam: PF01011; Bacterial_PQO; 7.
 DR PROSITE: PS00363; BACTERIAL_PQO_1; FALSE NEG.
 DR PROSITE: PS00364; BACTERIAL_PQO_2; FALSE NEG.
 KW Oxidoreductase; PQO; Quinate metabolism; Transmembrane.
 FT TRANSMEM 14 34
 FT TRANSMEM 41 61
 FT POTENTIAL.

FT TRANSMEM 68 88 POTENTIAL.
 FT TRANSMEM 90 110 POTENTIAL.
 FT TRANSMEM 127 147 POTENTIAL.
 SQ SEQUENCE 809 AA; 88196 MM; 71F67CEBA62BFCB CRC64;

Query Match 10.18; Score 313.5; DB 1; Length 809;
 Best Local Similarity 23.38; Pred. No. 2.2e-14;
 Matches 160; Conservative 80; Mismatches 262; Indels 185; Gaps 26;

3 PTSLMASAGALLAA--PAFAQYPTVDLALNPRAGEMTSTGQONENRHSPLTQIT 61
 148 PHEVTKASGEELPLVPDPKAKQY-----NMDDHNDGAGSREVALDQINR 193
 62 ENVSQDLQVW-----ARGMQP-KYQVTPPLIDGVYLANPGDVLOADAKTDLIME 113
 194 NNVSILKAMERFRTGDTTGTGNGCAEDQMPLQYKNVFLCTPHNNLFDALDADSGKQIMK 253
 114 HRRLPNATLINSFG--EPTRGALYGT-----NVEYV 144
 254 -----AEVNSTADAMERCGRVAFEDSTQPLVQPLAGATPAVALAANTCEPRRYTN 305
 145 SMDNHVALDTRAG-----QYTPVDNQGQEDMYSN-----SSGPIVANGVIAGSTC 192
 306 TVDRLTIANADTGARCKKDEGVNCTVNLHGLGENTKAPREVIASPTLGGTTIVGSRI 365
 193 QYSPF-----GCVSGHDSATGEELMRYNFIIPRAGEEDETWGNDEYARMWTGAMGQITTD 248
 366 ADVNAAADMPGGVIRAYVITGKLFMA--FDRPRNDPRVYALKPGIYKRSSTNSMAASTYD 423
 249 PVTPLVH--YGSTAVGASSETQRTPGCTLYGINTFAVRPDGELIYWHQTLPRDNMQ 306
 424 PQMTVFLPMGSSSDVWGGNR--TAADHKYNTSV-LALDATTGKEKWTYNTVHNDLMD- 479
 307 ECTPEMWTNVDPSTEMEGLOSINPNAATGERRVLTGVCRTGTMWQDAETGEFL-- 364
 480 -----FDLPMPQS-----LVDFPMKDGTTYKPAVYIGT--KSGQFVLDRTVYKPLIK 524
 365 -----WARTDNYNMIESIDENGIVYNE-----DALIKELD--VEYD 400
 525 VIEPIKVIADIPRGQYKTPRSVEMPQIQNTLKESDMGATPFQOLMRINPKSRID 584
 401 -----VCPITLGRHDPMSALNDSCITF-----IPL 427
 585 GLYTPAGTIVSLSPFGSISGMMGSIADFPTHRYMEFVNDRLGLMTIOLIKOTBEDIKIOA 644
 428 N---NVCYDMAADVDEFTSMQVNTSNVTKL--PRGKDMIGRIDALDISGRILMSYERA 482
 645 NGGKAVTGMGAIVPMKGTPTKVKKNRPMFSLGIPCCQPRPGTMTALDMKTRQYAWOVPLG 704
 483 AANYS-----PVLSTGGVLENGSTDRYFRALSOETGETLMQTRL 522
 705 TIQGTGPMGIKMGLKAPIGMTPTGGPMATGGLVFPAATQDYLLRANFSSNGEIMKARL 764
 523 ATVASQAISY--EVDGMQYVATAGG 547
 765 PVSGQTPMSYMSKPTGKYVYVVSAG 791

RESULT 15
 QUITA_XANCUJ STANDARD; PRT; 790 AA.
 AC Q9XD78;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PROBABLE QUINATE DEHYDROGENASE [PYRROLOQUINOLINE-QUINONE]
 DE (EC 1.1.99.25).
 GN QUMA.
 OS Xanthomonas campestris (pv. juglandis).
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 NC Xanthomonas.
 NCBI_TaxID=44291;

RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN-C5;
 RX MEDLINE=20063481; PubMed=10594704;
 RA Lee Y.-A., Lo Y.-C., Yu P.-P.;
 RT "A gene involved in quinate metabolism is specific to one DNA homology
 group of Xanthomonas campestris";
 RL J. Appl. Microbiol. 87:649-658(1999).
 CC -1- CATALYTIC ACTIVITY: QUINATE + PYRROLOQUINOLINE-QUINONE = 5-
 CC DEHYDROQUINATE + REDUCED PYRROLOQUINOLINE-QUINONE.
 CC -1- COFACTOR: PQO (BY SIMILARITY).
 CC -1- PATHWAY: QUINIC ACID CATABOLIC PATHWAY; FIRST STEP. THIS PATHWAY
 CC ALLOWS GROWTH OF BACTERIA WITH QUINATE BY ITS CONVERSION TO
 CC PYRROLOQUINOLINE AND SUBSEQUENT METABOLISM BY THE BETA-KETOADIPATE
 CC PATHWAY.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQO DEHYDROGENASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF109471; AAD38453.1; -
 CC Pfam: PF01011; Bacterial_PQO_6.
 CC DR PROSITE: PS00363; BACTERIAL_PQO_1; FALSE_NEG.
 CC DR PROSITE: PS00364; BACTERIAL_PQO_2; FALSE_NEG.
 CC KW Oxidoreductase; PQO; Quinate metabolism; Transmembrane.
 FT TRANSMEM 22 42 POTENTIAL.
 FT TRANSMEM 48 68 POTENTIAL.
 FT TRANSMEM 77 94 POTENTIAL.
 FT TRANSMEM 106 126 POTENTIAL.
 SQ SEQUENCE 790 AA; 82896 MM; B75F29B52A49EE6F CRC64;

Query Match 8.9%; Score 276; DB 1; Length 790;
 Best Local Similarity 22.1%; Pred. No. 8.7e-12;
 Matches 164; Conservative 79; Mismatches 262; Indels 236; Gaps 33;

9 ASAGALALLA--APAFQYPTVDE--LLANPRAG--EMISYQONENRHSPLTQI 59
 110 ATVALAVALVAGIGMEVPHRPVAGNAPGMAVPPGSOVQOANSAYGNDGSGRFAALDI 169
 60 TIENVGOLQVWARGQPKV-----QVTPILHDGVYLANPGDVLOADAKTDG 109
 170 NRSN-GRPAAGSPGPTTPEILANSNGAEDQLTPLQVGEKVFCTPHNNLIALDASTGK 228
 110 LIWEHRRLPNATLINSFCEPTRGALY-----GTN-----V 141
 229 QLM--RRE--INASSVWQRCRGICLPADAAALAPSVANSPILAAVYVAGANCRRRL 283
 142 YFVSWDNLVALDTAT-----GQYTFVDNQGQED--MVSNSGPIYANGVIAGS 190
 284 FNTIDGRLIADVADTGACQGFSGNGQVDLKAAGAAPDEPYQLTSPVLVAGTVVGR 343
 191 T-----COYSPRCFVSGHDSATGEELMRYNFIIPRAGEEDGTWKNDE-----A 235
 344 TRADDNVQTDMPGAVRG-----SMWSP--VRSAGLDP-----GNPHDROAPAAGSSIV 390
 236 RMMTGAMGQITVDPVYTNLVHGSTAVGPASFTQRTPGTLYGTN-----TRP 285
 391 RSTPNVWAPMSYDAAMNYF-----LPLGGPSTDLGARTALDHRYGASVIAL 439
 286 RPDGTGIYWRHOTLPRDNMQDECTFEMAVTVNDVQSTEMEGLOSINPNAATGERRVLTG 345
 440 DATGAEKWVQYTVHNDLMD-----FDLPMPQS-----IDPPNDGSHTPAVV 483
 346 VCKTKGTMMQPAEAGEFLW--ARDT-----NY-----QNMIESIDENGIVYNEAI 391
 484 IGTKAGQIVLDRAVATGKPLTEVEVVKGSIDIAHEQYAPTPQLSVGMPOIGTKHLETS 543

```

QY 392 LKELDEVYDVC-----PTFLGGRDMP$AALNPD$GIY--- 424
      : : |
Db 544 WCATAMDQMLCRIAERKQMRKEGLYAPGTDVLSFPGSLGGNMGGISTDPVHDVVFAND 603
      : : |
QY 425 -----IPLNN-----VCYDMMAVDOFT$MDVYNT$NVT$KL-----PPG 458
      : : |
Db 604 MRLGLWQMTIPADTRKAEAGGGEAVNTGMGAVPLKGTPIYAVNKNRFL$ALGIPCQARP- 662
      : : |
QY 459 KDMIGRIDAIDISTGRTLMSVERAAN-----YSPV-----ISTGGVLFN 499
      : : |
Db 663 ---YGTLSAIDIKTRSIAMQYVGTVDTPGPRGIMHLPPIGMPTLGGTLSTOGGLVFI 719
      : : |
QY 500 GGT-DRYEPAL$QETGELTMO$RLATV$AGALSY--EVDGMQYVATAGGVSYSGLNS 556
      : : |
Db 720 AGTQDYLERAFDSATGKELMKGRLPVGSOGPIYVSHKTKQYVVISAGGAR----- 772
      : : |
QY 557 ALAGERVDSTAIGNAVYFALP 578
      : : |
Db 773 -----QSPDRGDYVIA$LP 787
      : : |

```

Search completed: August 8, 2001, 19:44:11
 Job time: 183 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 2001, 19:40:38 ; Search time 27.61 Seconds
(without alignments)
2774.524 Million cell updates/sec

Title: US-08-934-506a-5

Sequence: 1 MKPSTLWASGALALLAAP.....GERVDSFAGNAVVFALPQ 579

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SPREMBL_16:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mammal:*
8: sp.mhc:*
9: sp.organelle:*
10: sp.phage:*
11: sp.plant:*
12: sp.spontent:*
13: sp.unclassified:*
14: sp.vertebrate:*
15: sp.virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	616	19.9	698	2	Q9KH03
2	583	18.9	708	2	Q46444
3	561	18.2	742	2	053362
4	560	18.1	601	2	09EYW8
5	549.5	17.8	695	2	Q9F902
6	537.5	17.4	601	2	P71509
7	479.5	15.5	633	2	024759
8	473	15.3	599	2	Q9L935
9	439.5	14.2	573	2	059540
10	365.5	11.8	790	2	09X255
11	327	10.6	803	2	Q9L115
12	324.5	10.5	786	2	P95466
13	281.5	9.1	644	2	Q52551
14	276.5	9.0	639	2	P77931
15	218	7.1	470	2	030326
16	208.5	6.7	182	2	032699
17	203.5	6.6	179	2	032697
18	200.5	6.5	180	2	032700
19	200.5	6.5	180	2	032696

20	200.5	6.5	181	2	032615	hyphomicrob
21	199.5	6.5	181	2	032692	hyphomicrob
22	197.5	6.4	181	2	032703	hyphomicrob
23	196.5	6.4	181	2	032621	hyphomicrob
24	195.5	6.3	182	2	032706	hyphomicrob
25	193.5	6.3	180	2	032701	hyphomicrob
26	191.5	6.2	182	2	032612	hyphomicrob
27	191.5	6.2	184	2	032693	hyphomicrob
28	190.5	6.2	172	2	033894	hyphomicrob
29	186.5	6.0	172	2	033882	hyphomicrob
30	186.5	6.0	179	2	032707	hyphomicrob
31	185.5	6.0	180	2	032704	hyphomicrob
32	185.5	6.0	180	2	032694	hyphomicrob
33	185	6.0	171	2	033884	hyphomicrob
34	183.5	5.9	179	2	032702	hyphomicrob
35	182.5	5.9	184	2	032695	hyphomicrob
36	181.5	5.9	185	2	0918K8	uncultured
37	174.5	5.6	179	2	032698	hyphomicrob
38	172.5	5.6	185	2	0918L0	uncultured
39	169.5	5.5	185	2	0918K9	uncultured
40	169	5.5	499	2	051842	051842 thioacillin
41	168.5	5.5	172	2	033889	033889 unidentified
42	168.5	5.5	172	2	09FDE8	09FDE8 environment
43	167.5	5.4	172	2	033886	033886 unidentified
44	167.5	5.4	407	1	027529	027529 methanobact
45	167.5	5.4	524	2	Q9PB85	Q9PB85 xyliella fas

ALIGNMENTS

RESULT 1
ID Q9KH03 PRELIMINARY: PRT; 698 AA.
AC Q9KH03;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE TERAHYDROFURFURYL ALCOHOL DEHYDROGENASE.
OS Alcaligenes eutrophus (Ralstonia eutropha).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=510;
RN [1]
RP SEQUENCE FROM N.A.
RA Zarrat G., Schraeder T., Andreesen J.;
RT "Molecular analysis of PQQ-dependent THFA-DR.;"
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF277373; AAF86335.1; -;
DR InterPro: IPR000345; -;
DR InterPro: IPR002372; -;
DR InterPro: IPR003088; -;
DR Pfam: PF000034; cytochrome c; 1.
DR Pfam: PF01011; Bacterial PQQ; 6.
DR PROSITE: PS00190; CYTOCHROME C; UNKNOWN_1.
DR PROSITE: PS00190; CYTOCHROME C; UNKNOWN_1.
SQ SEQUENCE 698 AA; 75471 MW; 9C87F180DB12FB2 CRC64;

Query Match 19.9%; Score 616; DB 2; Length 698;
Best Local Similarity 30.3%; Pred. No. 2.8e-35;
Matches 186; Conservative 93; Mismatches 247; Indels 88; Gaps 25;
QY 10 SAGALALIAAPFA--CVTPVDELLANPPAG--EWISYQONENRHSPLTOITTEWNG 65
DB 14 AAASVALPAPAFGANAAARVDGAIRANBAATPMPSTGLDYAEFTFSKLEQVNGNVR 73
QY 66 QLOLVARAGMPGK-VQVPLIHGVMYLANPGVIOATDAKTGDLMEHRQLP-NIAT 123
DB 74 NLGLAWSDESTRGVATPLVDGVYVSAPMSVVAIDAKTGKRLMTYDPOVPRDAY 133
QY 124 LNSGPEPTRGALVGTNVYFVSMNDHLVALDTATGQVTFD----VDRGQEDMVSNSGP 179
DB 134 KGCCDVYNRGVALYQGRVFGAFDGRVLVAIDAATGKRVWQDTLVDSKSYTI---TGAP 190

```

QY 180 IVANG-VIVAGSTQYSPGCEVSGHDSATGEELMRYEIP-----RAGEEDETW 229
DB 191 RYVNGVILGNGAEYGVNG-VITAYDAETGKQOMWYVYVPGDPAFENEMAKAAATV 249
QY 230 GNDYER-WMTG---AMQIYDPTNLVHGSTAVGASSETQGTGPG-PLYGTNRF 283
DB 250 --DPSGKWIINGGGGVWMTMAFDPELNTMYIGTGNAWSKRLSPKGGDNLVYASV-V 306
QY 284 AVRPDGEIVMRHQLPRDNMOCEFEEMVYN--VDVQPSDEMGLSINPNATGER 341
DB 307 ALMPDGETVMIHQETPGRNMDYTSQDILLDLKIDGP-----RK 348
QY 342 VILGVNCKTGTMMQDAETGEELMADTYNQMMIESIDENG--IYVNEITALKEIDVEY 399
DB 349 VILHAP-KNGFEFVIDRTNGKFLSAKNFVDVWASGYDNGRPVETPPADTSGKA----- 403
QY 400 DYCPTFLGRDMPSAALNDSGITFPLNNVCYDMAVDQETSMDYNTSN----- 451
DB 404 DVPVGPFGAHNMHSMFHPKGLAFIPQHPVL-TLADNKEW---VHNOKDSPAHRGV 458
QY 452 -----VTKLPCKMIGRIDAIDISTGRTLSVERAANSPVLSGGVLENGGTD 504
DB 459 GNLGMLVNAEPFRSKPEKRLTAMPDLQAKAVHHDHAGPMNGTLATAGMLVFGTADG 518
QY 505 YRRALSOETGETLMOTRLATVASGOAISYENDGMQYALA-GGVSYSGLNSALAGERV 563
DB 519 RLVAVHAATGKELMQAPGSSGVVAPVYLLDGRQYVSAVWGMGVYG-----LSQKGS 572
QY 564 DSTAIGNAVYVAL 577
DB 573 DRQAPGR-VYTEVL 585

RESULT 2
Q46444 PRELIMINARY; PRT. 708 AA.
AC 046444
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE OUNOHAEMOPROTEIN ETHANOL DEHYDROGENASE TYPE I PRECURSOR (EC 1.1.99.-)
  (OH-EDH1).
GN OHEDH.
OC Comamonas testosteroni (Pseudomonas testosteroni).
OS Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.
OX NCBI_TaxID=285;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15667;
RX MEDLINE=96184549; PubMed=8654419;
RA Stoorvogel J., Kraayveld D.E., Van Sluis C.A., Jongejan J.A.,
  De Vries S., Duine J.A.;
RT "Characterization of the gene encoding quinoxaloproduct ethanol
  dehydrogenase of Comamonas testosteroni."
RL Eur. J. Biochem. 235:690-698(1996).
RN [2]
RP SEQUENCE OF 32-54 AND 477-490.
RC STRAIN=ATCC 15667;
RX MEDLINE=95324580; PubMed=7601151;
RA De Jong G.A.H., Geerlot A., Stoorvogel J., Jongejan J.A., De Vries S.,
  Duine J.A.;
RT "Quinoxaloproduct ethanol dehydrogenase from Comamonas testosteroni.
  Purification, characterization, and reconstitution of the apoenzyme
  with pyrroloquinoline quinone analogues."
RL Eur. J. Biochem. 230:899-905(1996).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=86242113; PubMed=3521592;
RA Groen B.W., van Kleef M.A., Duine J.A.;
RT "Quinoxaloproduct ethanol dehydrogenase apoenzyme from Pseudomonas
  testosteroni."
RL Biochem. J. 234:611-615(1986).

```

```

CC -1- FUNCTION: CATALYZES THE DYE-LINKED OXIDATION OF PRIMARY ALCOHOLS
CC TO THE CORRESPONDING ALDEHYDES AND THE (SUBSEQUENT) OXIDATION OF
CC THE ALDEHYDES TO CARBOXYLIC ACIDS.
CC -1- COPFACTOR: POQ, HEME, AND CALCIUM.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: PERIPLASMIC (POTENTIAL).
CC -1- INDUCTION: BY ETHANOL AND BUTANOL.
CC -1- SIMILARITY: TO ALCOHOL DEHYDROGENASE (ACCEPTOR) PROTEINS.
DR EMBL: X81880; CA57464.1; -.
DR HSSP: P38539; 4AAB.
DR InterPro: IPR001479; -.
DR InterPro: IPR002372; -.
DR InterPro: IPR003088; -.
DR Pfam: PF00034; cytochrome_c; 1.
DR Pfam: PF01011; Bacterial_POQ; 6.
DR PRINTS: PR00605; CYTCHROMECIC.
DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
DR SIGNAL: POQ; Heme; Calcium; Oxidoreductase; Periplasmic.
FT SIGNAL 1 31
FT CHAIN 32 708
FT TYPE I.
FT BINDING 635 635
FT BINDING 638 638
FT METAL 639 639
FT SEQUENCE 708 AA; 76822 MW; 99AB54BDD6ACAB3 CRC64;

Query Match 18.9%; Score 583; DB 2; Length 708;
Best Local Similarity 27.9%; Pred.No. 6,1e-33;
Matches 173; Conservative 95; Mismatches 271; Indels 82; Gaps 22;

QY 3 PTLMSASAGALLAARAFAYQVTPV-----DELANPP-AGEWISYQNGEN 50
DB 12 PGRVWMLAACLG--SAARAFQTPGPAQAQAAVQYDVGFIIRANARTDMPIGVAAE 69
QY 51 YRHSPFLQITTEVNGOLVWARGMOPGK-VQYTPLIHGDVWVLANPQVIOALDKTDD 109
DB 70 TRSRDLQIINAAVKKDLGLMSTNLESTRGEVTPVVDGIMVSSWSVHAIDRTN 129
QY 110 LIWEHRRLPNIATLNSFGP-PTRGALYGTNYFVSWNHLVALDTATG-----QVTPD 163
DB 130 RIMTYDPQIDRSTGFGCCDVVNRGVALMKGYVYGAMPGRILALDAKGEVHQNTFE 189
QY 164 VDGQGGDMVNSSGPRTVANGVLAAGST-CQYSPGCEVSGHDSATGEELMRYEIP--- 219
DB 190 GQKGS-----LITGAPRVEGKVIIGKRAEYVGRG-YITAYDAETGERKRWFSVGPDP 244
QY 220 -----RAGEEDETWGNDEYEAHMTGA-----WQIYDPTNLVHGSTAVGASSETQ 268
DB 245 SKFEDESKMARATM--DPSGKWIINGGGGVWMTMAFDPELNTMYIGTGNAWSKRLSPKGGDNLVYASV 302
QY 269 RGPFGTLYGTNTRFVAVRDPDTEIVMRHQLPRDNMOCEFEEMVYNVDVQPSDEMGL 328
DB 303 RSPKGGDNLVYASVVALDPTGKWKHYQETPDNDYTSQDILLDLKIDGP----- 353
QY 329 QSLNPNAATGERVVLGVNCKTGTMMQDAETGEELMADTYNQMMIESIDENG-IYVYN 387
DB 354 -----AGKPRVILHAP-KNGFEFVIDRTNGKFLSAKNFVDVWASGYDNGRPVETPPADTSGKA 404
QY 388 EDALKEIDVEYVCPFTFLGRDMPSAALNDSGITFPLNNVCYDMAVDQ-ET----- 442
DB 405 --AAADSGKPDQAVGPRGALNNHMSNTPOTGLVYLPQAVVPAVNLDDKKHFEHQAGP 462
QY 443 ----SMDVYNTSNVTKL-PGKDMIGRIDAIDISTGRTLSVERAANSPVLSGGVLE 497
DB 463 GKQSGTGWNTAKFFNAEPBKSKPRGRLAMPVQAKAAVSEHVSPPWNGGLTLTAGNV 522
QY 498 FNGGTIRYRRALSOETGETLMOTRLATVASGOAISYENDGMQYALA-GGVSYSGLNS 556
DB 523 FQGTADGRLVAVHAATGKELMQAPGSSGVVAPVYLLDGRQYVSAVWGMGVYG----- 577

```

QY 557 ALAGEVDSTALGNAYVEAL 577
DB 578 -LAARATEROGPG-TYTFV 596

RESULT 3
ID 053362 PRELIMINARY; PRT; 742 AA.
AC 053362; 044159;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE DEHYDROGENASE SUBUNIT OF ALCOHOL DEHYDROGENASE.
OS Acetobacter pasteurianus.
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
OC Acetobacter.
OX NCBI_TaxID=438;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NC11380;
RX MEDLINE=94042848; PubMed=8226628;
RA Takemura H., Kondo K., Horinouchi S., Beppu T.;
RT "Induction by ethanol of alcohol dehydrogenase activity in Acetobacter
RT pasteurianus."
RL J. Bacteriol. 175:6857-6866(1993).
DR HSP; P38539; 4AAH.
DR InterPro; IPR001479; .
DR InterPro; IPR002372; .
DR Pfam; PF01011; Bacterial_P00; 6.
DR PROSITE; PS00363; BACTERIAL_P00_1; 1.
DR PROSITE; PS00364; BACTERIAL_P00_2; 1.
SQ SEQUENCE 742 AA; 81680 MW; 031C4E9B6B13EC98 CRC64;

Query Match 18.2%; Score 561; DB 2; Length 742;
Best Local Similarity 28.5%; Pred. No. 2.4e-31;
Matches 173; Conservative 90; Mismatches 248; Indels 96; Gaps 21;

QY 11 AALALLAFAAQTPTPTDEL--ANPPAGWISYGOENYRHSPLQITTEVVGOLQ 68
DB 23 AALPFAAFAADGAGNGEAIHADHPENWLSYGRYSQSPIDQIRSNVGLK 82
QY 69 LVMARGMOGKVO-VTPILHDGVMYLANPGVYQAIDAKTGDLIWEHRRLP-NIATLNS 126
DB 83 LAMVYTLIDNRQGEATPLVVDGIMATITWMSKEMALDADATGKLWQYPRKVGNIADGC 142
QY 127 FGEPTRGMALYGTNVYFVSMNDHLVALDTANGVTFD-----DRGGEEDMYSNSGPT- 180
DB 143 CDTVNRGAGYMGKVFMGTFDGRILVAADAKTGKRYWEVNTIPADASLKGORSYTVDAVR 202
QY 181 VANGVIVAGSTOQSPFGC--FVSGHDSATGELMRYNFIRPAGEGD----- 226
DB 203 VAKGVLILGN--GGSEFGARFVSAPDAETGKLKMFYTVPNKKPEPHADVANDVMSKA 260
QY 227 -ETWGDYERAMMT-----GANGQITTYDPTNLVHGSTAVGASERGTGPTGLY--- 277
DB 261 YKTWGP--KGAVYRGGGCTVWDSLYVDPSDLI---LAVG-----NGSPWNKKTSE 309
QY 278 --GTN---TRAVRDTGEIYVHQTLPDRNDQECTFEEMAVTVNDVOPSTEMEGLQSI 331
DB 310 GIGSNFLGSIVALKRETEGYVWHQATPMDQMDYTSVQOIMTLMPV----- 357
QY 332 NPRAAAGE--RYLVGVPCKTGMQFQDAETGEFLIARQNYNMESIDE--NGITVND 389
DB 358 ----NGEMRHVYIWAAP--KNGFFYVLDATKTEGLAGKNTVYONWANGDLPLTGRPIYND 411
QY 390 AILKEIDVEDVOCPTFLGGRDWPSSALNDSGIYFPLINVCY-----DMMK 436
DB 412 GLTTLGKFWYGIPLGLAHNFEMGAYSPKTHLVYLPAAHQIFPGYKNGVGGKPPHDAWN 471
QY 437 VDOEFTSMQVYNTSNVTKLPKGDMDIGRIDADIDISGRKLMSVERAANYSVLSTGGCV 496

DB 472 VGLDMTKNGLPDTPPE-ARTAYIKDLHGMLLAMDVYMETVWKIDHKGPWNGVYLATGDL 530
QY 497 LFNGGTDREYFRLASQETGETIMOTRLATVAGSQAISYEDVMQYVA-----IAGG 546
DB 531 LFOGLANGEEHAYDATNGSDLYKFDASGILASPMYTVNGKQYVAVEVWGGIYIISWG 590
QY 547 GVSYSYG 553
DB 591 GVGRTSG 597

RESULT 4
ID 09EYW8 PRELIMINARY; PRT; 601 AA.
AC 09EYW8;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE METHANOL DEHYDROGENASE LARGE SUBUNIT-LIKE PROTEIN.
GN MXAF.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=982;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RM1021;
RA Penner B.J., Tiwari R.P., Dilworth M.J.;
RT "Regulation of Cl assimilation in Sinorhizobium meliloti";
RT Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF09488; A631643.1; .
SQ SEQUENCE 601 AA; 65686 MW; 09F743726B2E2A83 CRC64;

Query Match 18.1%; Score 560; DB 2; Length 601;
Best Local Similarity 27.1%; Pred. No. 2e-31;
Matches 171; Conservative 109; Mismatches 248; Indels 104; Gaps 24;

QY 14 LALLAFAFAAQTPTPTDEL--LANPPAGWISYGOENYRHSPLQITTEVVGOLQ 71
DB 8 LAIMISGGAQVAFANDELQKLIDP--NQWAIQTGDYANRLYSKLQIDKNDYKGLQVAV 66
QY 72 --ANGMOGKVOYTPILHDGVMYLAN--GVYQAID--AKTGDLIWEHR--ROLNINATLNS 126
DB 67 TFSYGVLRGH--EGSPILVIGDLMVYTPFPMTVYALDSKQGLYKPEKQDPVIVPMC 125
QY 127 FGEPTRGMALYGTNVYFVSMNDHLVALDTANGVTFDVGQGEEDMYSNSGPT-VANGV 185
DB 126 CDTVNRGAGYVADNKIFLHQADTTYVALLDADATGKRYWVKNGDATTGNTATYVAVDKI 185
QY 186 IVAGSTOQSPFGCFVSGHDSATGELMRYNF-----PRAGEGDE 227
DB 186 LVGISGGEFVGRG--HYVAYSMDGKVLWRCYSMGPSDSDLIDPEKTHLKGPKGDSGLT 244
QY 228 TWGNDYERAMMTG---AMGQITTYDPTNLVHGSTAVGASERGTGPTGLYNTNRA 284
DB 245 TWEDD--QWMTIGGQITWGSYDPEENLYYGGNPSYMPQIR--PGDNRW-SMTIFA 296
QY 285 VRPDTGEIYVHQTLPDRNDQECTFEEMAVTVNDVOPSTEMEGLQSIINPAAAGEERVL 344
DB 299 RDVDTGAKKWLYQMTPHDEMVDGVNEMLITGQHLIDK-----DRKLIT 342
QY 345 GVPCKTGMQFQDAETGEFL-----WARTNYNMIE-----SIDNGI 383
DB 343 HFD--RNGFYMDRVTGELIAEKEYDPTVNATGEVMDPKSDKGRPOVYAOYSTEONG- 400
QY 384 VTVNEDALIKELDEYDVCPTFLGGRDWPSSALNDSGIYFPLINVCYDMMADVOEFTS 443
DB 401 ----EDNTT-----GCPALGTGKQDQPRAYSPTKTELEFVPPNHQCMDEPRPVSYTA 450
QY 444 MDVYNTSNVTKLPKGDV---IGRIDADISTGRKLMSVERAANYSVLSTGGVLENG 500
DB 451 GQPIYVATLSMYP--KDSHGGMGNFIACDNKEKIKNSLPEPFVSWGALATADVYFYG 509

QY 501 GTDYPFALSOETETLMOTRLATVASGOAISYEDMOQVATAGS-----GVSYGSLN 555
 Db 510 TLBGYLAKVADATKATKELRYKTPSGVIGANNMYTARBGKQVAVLSSGVGMAGIGLAAGLT 569
 QY 556 SALAG-----ERVDSIAIGNAVYVFPALPO 579
 Db 570 NPTEGLAVGVGSDLSNMYNLAGGTLTEFKLPE 601

RESULT 5
 Q9F902 PRELIMINARY: PRT: 695 AA.
 ID Q9F902; 01-MAR-2001 (Tremblrel, 16, Created)
 DT 01-MAR-2001 (Tremblrel, 16, Last sequence update)
 DT 01-MAR-2001 (Tremblrel, 16, Last annotation update)
 DE ALCOHOL DEHYDROGENASE.
 GN ADH.
 OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=316.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BC-2;
 RA Chang C.-H., Herrick J.B., Okinaka R.T., Brainard J.B.,
 RA Terrylliger T.C.;
 RT "Identification and characterization of genes activated by 2-
 RT chloroethanol in Pseudomonas stutzeri BC-2."
 RT Submitted (Aug-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF176640; AAG09249.1;
 DR EMBL: AF176640; AAG09249.1;
 SQ SEQUENCE 695 AA; 75842 MW; D29A698A642E1A83 CRC64;

Query Match 17.8%; Score 549.5; DB 2; Length 695;
 Best Local Similarity 29.2%; Pred. No. 1.4e-30;
 Matches 168; Conservative 89; Mismatches 259; Indels 59; Gaps 21;

QY 15 ALLAAFAFQVPTVDEL--ANPPAGEWISYGOQENVRHSPLOITTEVNGOLVMA 72
 Db 16 ALLVA-AGQAQAKVDDEAARASEQDSSELSHGRTYAEORSPLQIDAGNVGKIGLAWY 74
 QY 73 KGMQGRK-VQVTPLIHDGVMIYLANGVYIAQIDARTGDLMEHRQDP-NIATLNSFGE 130
 Db 75 LDLENNRGLEATPLVSDGLYASLSMSRYMAVDLSRKLQDFQVDRGHSRYTCCDAV 134
 QY 131 TFGMALYGTNNYFVSWDNLVALDPTAGQVTPDVDRGQEDMVNSGPIYANG-VIYAG 189
 Db 135 NRGVALMNGKTVYVGLDRLALDAKTGRELMSQTTDPARFYSITGAPRVKGVITGN 194
 QY 190 STQYSPFGCFVSHDSATGEELMNYFIIPRAGE-----EGDET--GNDYEARM 237
 Db 195 GGAEGVGVG-FPSAYDAETGKAMRFYTP--GPAOPIYEPBELAEAKTKKGDY--W 248
 QY 238 MTG-----AMGOITDPTVNLVHGSTAVGPASETQRTGGTLYGTNTRFAVRDPDTG 293
 Db 249 KLGGGGTVDMSACYPDELDTLYIGTGNGSPWNRREIRSPGDNLYLSIALRPDSKLL 308
 QY 294 WRHOTLPNDNDQECTFEMMTNVVQVSTEMEGQISINPNAAGGERVVLGVCKGTGM 353
 Db 309 WHYQTTPEETWDPTATQOITLATL-----ELDG-----KPRVLMQAP-KNGFF 351
 QY 354 WQDAETGEFLMARDTNNOMIESID--ENGIVYNEDAILKELDEYDVCPFTLGGRM 411
 Db 352 YVLDRTGETELLSAEKFGVTAKEKVDLATGPRVYPSGRYRKEQVVM--PSSFGAHNM 408
 QY 412 PSALNPSTGYTFPLNNVVCYDMAVDOEPTSMVYNT---SVATKLPPCKDM-GRID 466
 Db 409 HSMSPNPQTGLMYIPYQRIPIGVYNEGATFKKIDGLNTGTGFSPTHEIP--RDVSGALL 466
 QY 467 AIDISTGTMSVERAANYSVPLSTGGGVLFNGTDRYFALSOEGETELMOTRLATVA 526

Db 467 AMDPYRQREANRVPHSFYNGWGTLLSTAGNLVFGQTAGDSQLAHAYSADKQRLMSFAQTGI 526
 QY 527 SQQAISYEDMOQVAT-AG-GVSYGSLNSALA 559
 Db 527 VAAPISFSLDGEQYVAVMAGGAAPLIGDADALA 561

RESULT 6
 P71509 PRELIMINARY: PRT: 601 AA.
 ID P71509; 01-FEB-1997 (Tremblrel, 02, Created)
 DT 01-FEB-1997 (Tremblrel, 02, Last sequence update)
 DT 01-MAY-2000 (Tremblrel, 13, Last annotation update)
 DE METHANOL DEHYDROGENASE LARGE SUBUNIT HOMOLOG.
 GN MXP.
 OS Methylobacterium extorquens.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Methylobacterium group; Methylobacterium.
 OX NCBI_TaxID=408;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AMI;
 RX MEDLINE=97312011; Pubmed=9168622;
 RA Christoserdova L., Lidstrom M.E.;
 RT "Molecular and mutational analysis of a DNA region separating two
 RT methyloctrophy gene clusters in Methylobacterium extorquens AM1."
 RL Microbiology 143:0-0(0).
 DR HSSP; P38539; AAAH.
 DR InterPro: IPR002372;
 DR Pfam: PF01011; Bacterial_PQQ; 7.
 SQ SEQUENCE 601 AA; 64952 MW; 68FA5C7059CBB239 CRC64;

Query Match 17.4%; Score 537.5; DB 2; Length 601;
 Best Local Similarity 26.4%; Pred. No. 7.8e-30;
 Matches 166; Conservative 115; Mismatches 267; Indels 81; Gaps 23;

QY 1 MKPISLMSAGSALALLAFAFQVPTVDELANPPAGWISYGOQENVRHSPLOIT 60
 Db 1 MRAVHLALGGLAA--ASPLANESVLKG--VANP--ABQVLOTVDYANTRYSKLDQIN 54
 QY 61 TENYGOLOLV--ARGQPKQVQVTPLIHDGVMIYLANP--GDVLOAIDAKTG-DLMEHR- 115
 Db 55 ASNVKNLQVAMTSTGLRSH--EGSPLVYGNINMYHTPPENIYALDLOGAATVKKYEP 113
 QY 116 RQLEPNATLNSFGEPTRGMALYGTNNYFVSWDNLVALDPTAGQVTPDVDRGQEDMVNS 175
 Db 114 KQDSVPIPVACCPTVNRGLAYADGAILLHQADTTLVSLDAKSGKVMYSKNGDPSKGETN 173
 QY 176 SSGIYANGYIVAGSTQYSPFGCFVSHDSATGEELMNYFI----- 218
 Db 174 TATVLPVKDRIYVIGISGEGFVQCHVATYADLKSGRKWRGYSIGPDQOLIVPEKTTSLG 233
 QY 219 -PRAGEEDETWQNDYEARMGTG---AMGOITDPTVNLVHGSTAVGPASETQRTGG 274
 Db 234 KPLAKDSLSLKTWBGD---QMKTGCGCTGWGFSYDPLDLATMYGSG--NSTGNPKORPBD 288
 QY 275 TLVGTNTRFAVRDPDTGSIWRHOTLPNDNDQECTFEMMTNVVQVSTEMEGQISINPN 334
 Db 289 NKW-SMTIMARNDDTGMKAVVYQMTPHDEWDPDGINEMILTD-----OKPDG----- 334
 QY 335 AANGERVVLGVCKGTMMQPDATGEFLMARDTN-YOMIESID-ENGIVYNEDAIL 392
 Db 335 ---KDRPLTHPD-RNGFGTTLDRATGEVLAKEFDPVVMATKVDLDGSKSTYGRPLV 390
 QY 393 KELDEYD-----VCPFTLGGRDPSAALNPDGSIYFPLNNVVCYDMAVDOEPTSM 444
 Db 391 SKSTEDNGSDVMSKICCPAALGTCKQDPAAPSFKGLGYVPTNNHVCMYDEFPRTYITG 450
 QY 445 DVYNTSVNTKLP-PGR-DMIGRIDALIDISTGRTLMSVERAANYSVPLSTGGGVLFNGGT 502

Db 451 QPYGATLSMYPAPSGHGMGFIAMDNLOGKIKSNPEQFSANGALATSGDLYFYCTL 510
 QY 503 DRYFRALSGEETLMQTRLATVASGOAISTEVDMQYVAIAG-----GVSYSGSLNSA 557
 Db 511 EGFILKAVDSKTKGELKFKPSPGIIIGNMYIEHKQKHVAVLSGVGMAGIGLAAGLTP 570
 QY 558 LAG-----ERVDSIAIGNAVYVFLP 578
 Db 571 NAGLGAGGYAALSSYTNLGGQLVFEFLP 599

RESULT 7

024759 PRELIMINARY; PRT; 633 AA.
 AC 024759;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE METHANOL DEHYDROGENASE ALPHA UNIT PRECURSOR (EC 1.1.99.8).
 GN MXAF.
 OS Hyphomicrobium methylovorum.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Hyphomicrobium group; Hyphomicrobium.
 OX NCBI_TaxID=84;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GM2;
 RX MEDLINE=97457202; PubMed=9311140;
 RA Tanaka Y., Yoshida T., Watanabe K., Izumi Y., Mitsunaga T.;
 RT "Cloning and analysis of methanol oxidation genes in the methylotroph
 Hyphomicrobium methylovorum GM2";
 RT FEMS Microbiol. Lett. 154:397-401(1997).
 RL EMBL, AB004097; BAA23272.1;
 DR HSSP; P38539; 4AAH.
 DR InterPro; IPR001479;
 DR InterPro; IPR002372;
 DR Pfam; PF01011; Bacterial_PQO_7.
 DR PROSITE; PS00363; BACTERIAL_PQO_1;
 DR PROSITE; PS00364; BACTERIAL_PQO_2;
 KW Signal; Oxidoreductase.
 FT SIGNAL 1 34
 FT CHAIN 35 633 POTENTIAL.
 SO SEQUENCE 633 AA; 69852 MW; BA7A23A2779E2C1B CRC64;

Query Match 15.5%; Score 479.5; DB 2; Length 633;
 Best Local Similarity 25.0%; Pred. No. 1e-25;
 Matches 158; Conservative 116; Mismatches 270; Indels 87; Gaps 22;

QY 6 LMASGALALLAAPAFQVPTVTELLANPAGE-WISYGNOMRHSPLTQITTEV 64
 Db 12 LMSASGMAIAVALLOGVASSAVANDKLELSKSNENWVPGKNYSNNYSESTQVNAENV 71
 QY 65 GOLQVWARGMOPKGV---QVTPPLIHIDGVYMLAN--PGDVIQAIIDAKTGDLIWEHR-RQL 118
 Db 72 KQKHAWS--FSTGELHGEAPLVIGDMYVHSSFPNKFALINDEGHILMOHSPKOD 129
 QY 119 PNATITNSGEPFRGMALYGTN-----YFVSMNHIALDTATGQVTFVDVDRGGEDM 172
 Db 130 PAARSVACCDLVNRGLAYVPGDKTPALVITKQLDGHVALNAKGEERKVENGDIKG 189
 QY 173 VNSSGPIVANGVTVAGSTQYSPFGCFVSGHDSATGEELMRYFI----- 218
 Db 190 QITLQAPYVYHDLAIYSSGAGELGVNGHTAVNVKTGEQAMRYVATGPEDEIGLADDFNS 249
 QY 219 --PRAGEE--GDETWNDEYARMWGA---WGQITVDPVTNLVHGSTAVGASERQRT 271
 Db 250 ANHYQCKGIGTATWEGD---AWKIGGTNMGYAYDPANLITYGSGNPAPWNETMR-- 304
 QY 272 PGGTIGTGNFAVVPDGTGFIWRHQTLPNDNMDDCEFTFEMMTVNDVQSTEMEGLOSI 331
 Db 305 PGDNKM-TMTITRADDTGKMKFGYOKTPEHDEWDFAGVAVIMLSE-----QIDKE----- 353

QY 332 NPNAATGERRVLTGVCCKTGIMQPDATGEELMA---RDTNYKNMIESIDENGIYVNE 388
 Db 354 -----GKKRKLTLTHDRNGIYVTLIDRENGDLISADKLDDT--VNVFKHVDLKSGLPVD 405
 QY 389 DALKELDVE-YVCCPTFLGGRWPSAALNPDGSIFFILNNVCYDMAVAVDEFTSMDY 447
 Db 406 PERGTBRDHKGTEICPSAMGYHNOGHDSYPTKQLFFKGINICHMDEFFMLPYRAGQFF 465
 QY 448 NTSNVTKLTP-PGKDM-----IGRIDAIDISTGRTLSVERAANYSPVLSTGGVLFNG 501
 Db 466 VGATLMWYEPKGRDRONYLGLGQIKAVNATITNYKKEHHERSVWGCLATAGNLVFGYT 525
 QY 502 TDYFRALSGEETLMQTRLATVASGOAISTEVDMQYVAIAG-----GVSYSGSLNS 556
 Db 526 LDGFLKARNSDTGELLMKHKLPSGIVIGYPTIEHKQGVYIAVMSGVGMGPVGLVDELDD 585
 QY 557 ALAG-----ERVDSIAIGNAVYVFLP 577
 Db 586 PTAGLGAVGAFKMLQRY--TOMGGSLEVFSL 614

RESULT 8

09L935 PRELIMINARY; PRT; 599 AA.
 AC 09L935;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE METHANOL DEHYDROGENASE LARGE SUBUNIT.
 GN MXAF.
 OS Methylovorus sp. (strain SSI / DSM 11726).
 OC Bacteria; Proteobacteria; beta subdivision; Methylophilus group;
 OC Methylovorus.
 OX NCBI_TaxID=81683;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SSI;
 RA Kim Y.M.;
 RT "Cloning and nucleotide sequence of mxaf gene of Methylovorus sp.
 strain SSI DSM11726";
 RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RL EMBL, AF184915; AAD56237.2;
 DR EMBL; AF184915; AAD56237.2;
 DR InterPro; IPR001479;
 DR InterPro; IPR002372;
 DR Pfam; PF01011; Bacterial_PQO_7.
 DR PROSITE; PS00364; BACTERIAL_PQO_2;
 SO SEQUENCE 599 AA; 65133 MW; DBF6F4B5D871BC91 CRC64;

Query Match 15.3%; Score 473; DB 2; Length 599;
 Best Local Similarity 25.2%; Pred. No. 2.8e-25;
 Matches 157; Conservative 108; Mismatches 281; Indels 76; Gaps 21;

QY 4 TSLMASGALALLAAPAFQVPTVTELLANPAGE-WISYGNOMRHSPLTQITTEV 63
 Db 6 TATGFAVAGLALSMBLBSVAADSLLEALGNP--NNMTQGTGYTQOHRSLSQITGN 63
 QY 64 VGOLQV---VWARGMOPKGVTPPLIHIDGVYMLAN--PGDVIQAIIDAKTGDLIWEHR-RQ 117
 Db 64 VKNLESGLGHSTGLTNGH--EGAPLVIGDMYVIMTPFPNNTFALINAEPEKIVQHRPKQ 122
 QY 118 LPNATITNSFEPFRGMALYGTN---YFVSMNHIALDTATGQVTFVDVDRGGEDMVSNS 177
 Db 123 DASVAAVACCDIVNRGLAYDDGHIFKTQLDGHVLAASDAKTGKELMKMENCDDPAVGSTITQ 182
 QY 178 GPIVANGVTVAGSTQYSPFGC--FVSGHDSATGEELMRYFI-----P 219
 Db 183 APFAVKGAVLVG--CSGELGVRGVTVAYDQKTGEIVWRSRATGPDELNLAKDFNKNP 240
 QY 220 RAGEE--GDETWNDEYARMWGA---WGQITVDPVTNLVHGSTAVGASERQRT 274
 Db 241 HYQWGLGTKTWEGE---GMRIGGTNMGYAYADPKMLFYFGSGNPAPWNETMR--PGD 295

QY	275	TLGVTTPRAVPEPDGEIYVWRHQTLPKPNMPOECEEEMVNTVNOVQSTEMEGLOSTNP	335
Db	296	NKM-TFTIARVDYDGAANKWQKQTPRHEBMDPAGYNQMLTDQAVNGKTO-----	344
QY	335	AATGERRVLTVPCVKTGTWMPQDAETGEFLAARDTN-YQNMIESIDENGCIYTVNEDAILK	393
Db	345	-----PLTTTHVD-ENGVIYTLNRTQSGISVQAAKVDAVAVNEFKKVDLKTGLPVRDPDEFST	397
QY	394	ELDVE-YDVCPTFLGGRBMSALNPDSGIYFIPILANNVCYMAAVDDEFLSMDEVYNTSNV	452
Db	398	RMDHKGTNLCPSAMGPFHNOGLDADVPDSRTYFEGLNHCMDMEPFMLPYRAGQFEVQATL	457
QY	453	TKLP-----PGKMDIGRIDALDISTGRTLSVERAANSPVLSYGGVTLFNGGTDYRFA	508
Db	458	AMYPGNGPTKKEMGOVLAMDSVGEVWTKMEKFPWSWGGLTAKGSLVYTNILDGIRKA	517
QY	509	LSQETGELTMQTRLATVVASGQAISETVDGMQYVAI---AGS-GVSYGSGLSNALG--	560
Db	518	LDKTKGKELMKKKMPSGAIGAPMSYAKKQYIATNYGVGMPGVGLVFLDTPSAGLGA	577
QY	561	-----ERVDSFAIGNAVYVFL 577	
Db	578	VGAFFKELQNTYTMGGGVAVFGL 599	
RESULT	9		
ID	059540		
AC	059540:	PRELIMINARY:	PRT: 573 AA.
DT	01-NOV-1996	(TREMBlrel. 01, Created)	
DT	01-NOV-1996	(TREMBlrel. 01, Last sequence update)	
DT	01-MAY-2000	(TREMBlrel. 13, Last annotation update)	
DE	METANOL, DEHYDROGENASE HEAVY SUBUNIT (EC 1.1.99.8) (ALCOHOL		
DE	DEHYDROGENASE (ACCEPTOR)).		
OS	Methylophilus methylotrophus (Bacterium W3A1).		
OC	Bacteria; Proteobacteria; beta subdivision; Methylophilus group;		
CC	Methylophilus.		
OX	NCBI_TaxId=17;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=W3A1;		
RA	Xia Z.X., Dai W.W., Zhang Y.F., He Y.N., White S.A., Boyd G.D.,		
RA	Mathews F.S.;		
RL	Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.		
CC	-I- CATALYTIC ACTIVITY: A PRIMARY ALCOHOL + ACCEPTOR = AN ALDEHYDE +		
CC	REDUCED ACCEPTOR.		
CC	-I- COFACTOR: POO.		
DR	EMBL; U01040; AAA83765.1; -.		
DR	HSSP; P38539; 4AAH.		
DR	InterPro; IPR001479; -.		
DR	InterPro; IPR002372; -.		
DR	Pfam; PF01011; Bacterial_POO; 7.		
DR	PROSITE; PS00364; BACTERIAL_POO_2; 1.		
KW	Oxidoreductase.		
SO	SEQUENCE 573 AA; 62635 MW; A06C9B3091BB8F0C CRC64;		

	Query Match	14.2%	Score 439.5;	DB 2,	Length 573;
	Best Local Similarity	24.5%;	Pred. No. 6e-23;		
	Matches	146;	Conservative	99;	Mismatches 259; Indels 93; Gaps 21.
QY	38	AGEMISYGQENQEVHRSPLEQTITTNVGOLOLW--ARGMGGKYVTPLIHGVMWL--	93		
		: : : : : :			
Db	13	AGAMPATSGSYGSQHNSPPLAQINKSNVKNKRAMSFSTGVNLGH--EAPPLYIGDMNVHS	71		
QY	94	ANPGDYIQALDAKTGDLLWEHR-RQLPNIALTNSFGEPTRMALYGTNYVFSMDNHLVA	152		
		: : : : : : :			
Db	72	APENNITYALINLNDPGKITWQHKKPODASTAKVAMCCDDVDYDRGLAYGAQGIYKKQANGHLLA	131		
QY	153	LDPATGQVTFEDPDROGGEDMYNSSGPIYANGVIYA GSTCOYSFFGC--FVSGCHDATGE	210		
		: : : : : : :			
Db	132	LDAKTKIRIMVEVCDPKKGFSLTPADPAVDYTLMG--CSGAEIAGRVAVNAFDLKTRE	189		

```

0Y 211 ELMR-----NYEIPRAE--EGEETGNDVEARMMGA---MGQITYP 249
Db 190 LKMRPAFGSDSVLADPFSANPHYGQFGLKRTKEGD---AMKGGGTMMGWATYP 246
QY 250 VTNLVHGSTAVGASAPETQGTGG---TLVGTNTRFAVRPDGTGEIVMRHOTLPEDNM 304
Db 247 KLTNEYFSGNBPAPNETMR--PGDNKWTMTIMGRDL-----DTGMAKKGYOKTPHDEM 298
QY 305 DQCEFFEMAMVNVQVPESTEMEGLOSTINPMATBERVLGVPECKTITMOPAEFGEFL 364
Db 299 DPAGYNQVLTID-----QPVN-----GKMPPLSHIDRNLIYTLUREGNLI 341
QY 365 WARDIN-YONNIESIDENGIVTNEDAILKELDVE-YDVCPTFLGDRMDESALNDPSGI 422
Db 342 VAEKEDPAVNVFKKYDLKTGPVADPEFAIRMDHKGNITCPASMGFINQGVDSIDESPRT 401
QY 423 YFPIINVQCYDMAVDOEFTSMDYNTSNVTKLP---PGKDMIGRIDAIIDISTGRTLWS 478
Db 402 LVAGINHTICMDWEPFMLPYRAGQFEGVATLAMPVGPNGPTKKEMGQIRAPDLTGKAKMT 461
QY 479 VERAAANTSPVLSTGGVLEFMGCDIRFRALSOETGTTLMQTRLATVYASQAISTEYDGM 538
Db 462 KWEFFAALGGTLTYRKGGIWMYATLDGLKALDNKDKEINFMFMPSGISGSPTYSFYFK 521
QY 539 QYVALAGVYSGSG-----LMSALG-----ERDSTALGANYVYFAL 577
Db 522 QYT-----GSMYTGVMPCVGLVFDLTDPSAGICAGCAFEELONHTNOMGGGLWVFSL 573

```

RESULT	10		
Q9XZS5			
ID	Q9XZS5	PRELIMINARY;	PRT; 790 AA.
AC	Q9XZS5;		
DT	01-NOV-1999 (TrEMBLrel. 12,		Created)
DT	01-NOV-1999 (TrEMBLrel. 12,		Last sequence update)
DT	01-MAY-2000 (TrEMBLrel. 13,		Last annotation update)
DE	GLUCOSE DEHYDROGENASE.		
CN	GDBB.		
OS	Patloea citrea.		
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;		
OC	Patloea.		
OX	NCBI_TaxID=53336;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=1056R.		
RX	MEDLINE=99303331; PubMed=10376838;		
RA	Pujol C.J., Kado C.I.;		
RT	"gdbb, a gene encoding a second quinoprotein glucose dehydrogenase in		
RL	Patloea citrea, is required for pink disease of pineapple.";		
EMBL	AF050503; AF050503; AAD23735.1; -.		
DR	InterPro: IPR001479; -.		
DR	InterPro: IPR002372; -.		
DR	Pfam: PF00101; Bacterial_POG_7.		
DR	PROSITE: PS00364; BACTERIAL_POG_2, 1.		
SQ	SEQUENCE 790 AA: 85681 MW: 888B7625E0A5F089 CRC64:		

	Query Match	11.8%	Score 365.5;	DB 2,	Length 790;	
	Best Local Similarity	23.9%;	Pred. No. 1.6e-17;			
	Matches 169; Conservative	78;	Mismatches 238;	Indels 221;	Gaps 34;	
QY	27	PVYDELLANPAGEMISYGONENRHSPLQLITTEWVGOLUWA-----RGMQPKG	79			
	: : : : : : : : :					
Db	152	PAHQPYSAK---DMTATGCTGAGGRYSLSQIMESYNKLNQVAMTYQSDDVKRRPPDPE	208			
QY	80	V--QVTPLIHDSVMFLANPGDVIAIDAKTDGLIMEHRROLPNIALINSFSEPT-RGMAL	136			
	: : : : : : : : : : :					
Db	209	ITNESPTLKVGLMLTYCTAHQILVAIDAATGEKM--REDPRKIKDPTFGNHTCGSVSY	265			
QY	137	YSGN-----VFVSNDNLHVALDLPATGCQTFFVDVRGGEDM-----	172			
	: : : : : : : : : : :					
Db	266	HGFNTADPDQAPSPASCSRRIATPVWDGGLAVLDENQSLDKDGNGGELDLOHQDN	325			

QY	173	---	VNSSSSIIIVANGVIA-AGS-TCOYS-PPGCVSHSDATGELRRNFIFIRP	222
Db	326	AFPGYEPTSPVITDKYIYIAGAVTDNYSTIRESG-VIRGFVDKTKGLLMA--FDGAK	382	
QY	223	E-----EGDETWGNDYEAARMGTGAMCOITVDPTNLVHYGSTAVGPASETQRCPTGL	276	
Db	383	DPNAIPADGQHCASANS-----PNSMAFAAADPKLIDLIVLEPMGYTTPIDIMGNGRTPEQEH	436	
QY	277	YGTNTRFAVVRDPTGEIWRHOTLPDRDMMDOCEFTPEMVTNVDOVPTSTEMEGIOSINPAA	336	
Db	437	YASGI-LALNATTTGKEAFQYOTVHHDLMDMD-----IPAQPS-----LADITDK--	479	
QY	337	TGEREVLIVGVECKTGTMMQFPAETGEF-----	363	
Db	480	DGNKPVLIYVETKGNIFVLNRATGQVLVPAPERAVPOGAKGDHSTATQFPPELTFRPE	539	
QY	364	-----IMADITYNQNNI-----ESIDENQIVT-VNEDALILKELDVEDVCPFLGGRD	410	
Db	540	AKLIDSKKMWG-GTMDQDQICIRKRLRYNGTFPPPSQGL-----VFPGLNGFE	590	
QY	411	WPSAALNDPSGIYFIPLNVCYDMMAVDQETSMQVYNTFSNVTLLPKPKDMIGR-----	464	
Db	591	WGGISVN--SGSEFAIAN-----PMQLPVSFKLIPRPN-NMPEPKNAGSGSGEAG	640	
QY	465	-----IDAIDISTGRTLMSVEKAANYS-----	487	
Db	641	LOHMGVFPGVELNPFLLPCMEPWSGVSAVNLTNQIVMKKRTIGTIRDSGPTPL	700	
QY	488	-PV-----ISTGGGVLENGGT-DRYFPAISOETGETILMOTRIATVASGQ--ISY	533	
Db	701	PTITGTGPMLGAPITTAAGNIIEFIATLTDDYIRAVSVRQGLKLMOARLP--AGGQATPMTY	758	
QY	534	EVDGMOY-VALAGVSVYSGSLNALSALAGERDSTVANNVYFALP	578	
Db	759	EVDGKQYLIYVAGHGSGF-----TKMGDIITAYKLP	790	
RESULT	11			
	091115	PRELIMINARY:	PRT:	803 AA.
AC	091115			
DT	01-MAR-2001	(TREMBLrel. 16, Created)		
DT	01-MAR-2001	(TREMBLrel. 16, Last sequence update)		
DT	01-MAR-2001	(TREMBLrel. 16, Last annotation update)		
DE	GLUCOSE DEHYDROGENASE.			
GN	GCD OR PA2290.			
OS	Pseudomonas aeruginosa.			
OC	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;			
OC	Pseudomonas			
OX	NCBI_TaxID:287;			
OX	NCBI_TaxID:287;			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=PA01;			
RX	MEDLINE=204373737; PubMed=10984043;			
RA	Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,			
RA	Hickey M.J., Brinkman F.S.L., Huffrage W.O., Kowalik D.J., Lagrou M.,			
RA	Garber R.L., Collier L., Tolentino E., Westbrook-Webman S., Yuan Y.,			
RA	Brody L.L., Coulter S.N., Folger K.K., Kas A., Labdig K., Lim R.M.,			
RA	Reiter J., Sater M.H., Wong G.K.-S., Wu Z., Paulsen I.T.,			
RT	Reiser J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RT	*Complete genome sequence of Pseudomonas aeruginosa PA01, an			
RT	opportunistic pathogen.*;			
RL	Nature 406:959-964(2000).			
RL	EMBL: AE004654; AAC05678.1; -.			
DR	InterPro: IPR002372; -.			
PFAM	PF01011; Bacterial_P00; 7.			
SEQUENCE	803 AA; 86217 MW; 7DDF46B0FE5ED249 CRC64;			

Query Match	10.6%;	Score 327;	DB 2;	Length 803;
Best Local Similarity	21.9%;	Pred. No. 8.7e-15;		
Matches 166;	Conservative 84;	Mismatches 265;	Indels 244;	Gaps 32

OY	6	LMW-----	ASGALALAAAPFAQVTFVTDELLAN-----	36
Db	103	LLMLPWFRRPLLADGPAPLPGTAAVLGAAVAAGSG-FTINPQIYGRIDRDSGMTST		161
OY	37	-----PAEWISTYQONENYRHSPLTQITTEENVGQOLVWARGMOPGVK-----	Q 81	
Db	162	APAPDSDMWAYARTFEGRGYSPLKITPANGQGLEAMR--IRTGDPPTADDPLELTNE		219
OY	82	VTPILHDVGLANPGDVIOALDAIKGDLIMEHRROLPIRIATISFGEPT--RGAAIKGTN		140
Db	220	NTPLKVNGMLYACTAHSKVLADPDGTGAIMRFDQIOSPVGFKFAHMTGRGVSYDEE		279
OY	141	VY-----	FVSMNHLYALDPTATGQV-----TFDV 164	
Db	280	QYARSDVGAPRALLSERGAKAVASCPRLFLPTADARLAIINADNGKCEDPYKGAVDL		339
OY	165	DRGGE-----DMVSNSSGPIVANGVIVAG-----STCOYSPGECFVSGHSDATGELMRYN		216
Db	340	TAGIGPPTPGGYSTSPAAVTRLVIIIGHVADNESTNEPSS-VIRADVDHDKLW--NW		397
OY	217	FIRPAGEECGCTEMGNDYEAMRTGAMGQITTPYTNLVHGSTANGAPASEORG---TPG		273
Db	398	DSGNPDETEPLARKEY--TNSPNNMSTLASVBERLQGVY--LPLGNOMPQMGNRTPG		453
OY	274	GLTGTMTREFVRPDITGEIYWRHQTLPRNMDOECTEFEMVNTNDVOPSTEMEGLSINP		333
Db	454	AERKSAGI-VALLDNTCKLKNYQFTHHDLMDMD-----VGSQPTL-----LDL		496
OY	334	NAATGERRVLTVPCKTGTIMQDAETG-----		361
Db	497	KTAGGVPRALI-APTCKGSLYVLDRRDGPPIVPIREVPAPQAGVGDHTAPTQARSDNL		555
OY	362	-----EFLMA-----BDTNQNNITESIDENGIVTVNEDAILKELDEYDV		401
Db	556	LRPLTERDMWSSPEPDQMLCRIOFNSLRTEGQYTPPSQGSL-----I		599
OY	402	CPTLEGGDMPSALINDSGIYEIPLNNVCYDMAVADOET---SMDVYNFSNVTK-----		454
Db	600	YFGNVGVFNNMGCVSVDPVRQILFTSPRYMAFVSQVPRDKVPSGKREGETSGVQPTGA		659
OY	455	-----LPPGKMIGRIDALIDISTGRITLWSTERAAA-NISPY-----		489
Db	660	PYAVIMHPFMSPLPCQAPSWGCVAGIDIDTTAKVYWMQKNGTSRNPTRPVIGLTVGPS		719
OY	490	-----LSTGGAGVLENGGF-DYRFRALQSETGETLQOTRLAFVASGQA--ISYE-VDMQGV		541
Db	720	MGGSTTTAGAGVALISGLDQYLAHYDKKQKQMLQARLP--ACGQATPRSYTGKGRQYV		777
OY	542	AI-AGGGVSYSGINSALAGERVDSITAIGNAVYVFPQ 579		
Db	778	LIVAGGHSFG-----FMGDYIATVALPR 802		
RESULT	12			
P95466		PRELIMINARY;	PRT; 786 AA.	
AC	P95466:			
DT	01-MAY-1997 (TReMBLrel. 03, Created)			
DT	01-MAY-1997 (TReMBLrel. 03, Last sequence update)			
DT	01-JUN-2000 (TReMBLrel. 14, Last annotation update)			
DE	GLUCOSE DEHYDROGENASE [PYRROLOQUINOLINE-QUINONE] (EC 1.1.99.17).			
GN	GDH.			
OS	Pantoea citrea.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Pantoea.			
OX	NCBI_TaxID=53336;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND ACTIVITY.			
RC	STRAIN=1056R;			
RX	MEDLINE=97133947; PubMed=8979341;			
RA	Cha J.-S., Pujol C., Kado C.I.;			
RT	"Identification and characterization of a Pantoea citrea gene encoding			

RT glucose dehydrogenase that is essential for causing pink disease of
RT pineapple.
RL Appl. Environ. Microbiol. 63:71-76(1997).
CC -1- CATALYTIC ACTIVITY: D-GLUCOSE + ACCEPTOR = D-GLUCONO-1,5-LACTONE +
CC REDUCED ACCEPTOR.
CC -1- CORRECTOR: POQ.
CC -1- SUBCELLULAR LOCATION: PERIPLASMIC (BY SIMILARITY).
CC -1- MISCELLANEOUS: RESPONSIBLE FOR PINK TO BROWN DISCOLORATIONS OF
CC INFECTED PINEAPPLES UPON HEATING DURING THE CANNING PROCEDURE.
CC INFECTED BATCHES HAVE TO BE DISCARDED.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
DR EMBL: X85965; CA65229.1;
DR InterPro: IPR001479;
DR InterPro: IPR002372;
DR Pfam: PF01011; Bacterial_POQ_6.
DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
KM Oxidoreductase; POQ; Transmembrane; Periplasmic.
FT DOMAIN 1 13 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 14 55 POTENTIAL.
FT DOMAIN 56 59 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 60 80 POTENTIAL.
FT DOMAIN 81 85 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 86 106 POTENTIAL.
FT DOMAIN 107 115 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 116 136 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 137 431 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 432 452 POTENTIAL.
FT DOMAIN 453 652 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 653 673 POTENTIAL.
FT DOMAIN 674 700 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 701 721 POTENTIAL.
FT DOMAIN 722 786 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 724 770 POTENTIAL.
FT ACT SITE 462 462 BY SIMILARITY.
SQ SEQUENCE 786 AA; 86038 MW; 372402AAD8B67CC CRC64;

Query Match 10.5%; Score 324.5; DB 2; Length 786;
Best Local Similarity 22.7%; Pred. No. 1.3e-14;
Matches 170; Conservative 91; Mismatches 240; Indels 247; Gaps 36;

QY 2 KFTSLMASAGALLAALAPAFQAQ-----VTPYDELLANPP-----AGEMISYQGNQ 48
DB 115 KFLAMVIGLGINALLLGLASLHDPQETNGVLYNSDK-----PPASASASADWPAYGRQ 170
QY 49 EYRHSPLTOITTEVNGQLQVW-----ARGMQPGAV--QVPLPHDGVYLANPQDV 99
DB 171 EYVRSPLTOINDKNVQOLQVAMQFHTGDKHTANDPGEITNEVPLKYNMXYLCTPHQI 230
QY 100 IQAIDAKTGDLIWEHRROLPNIAI-----LNSFGEPTRGALYGVWYFVSW 146
DB 231 LIALDAAGREKWRFPDLPKSDPTFQHTTCRGVSHETKSVOGSSAPAACSRIFLPVD 290
QY 147 DNHLVALDTATGQVTFVDVROGQEDMSN-----SSGPYIANG--VIVAGSTCQ 193
DB 291 DGRLEFVALVALTGORSNF--ANNGELINLOHLPNAYPGGEPTSPILITDKVVIAGSVTD 349
QY 194 -----YSPGCVSGHDSATGEIMRNYPFRAGE-----EGDETNGWDYAKRMKTGAWG 243
DB 350 NLSIREPSG--VIRGFDIDSGKLLM--VEDPGAKDPNAPVAGQGFVANS-----PNSMA 400
QY 244 QITVDPTNLVHYGSTAVGPASSETORGTPGGLYGTNTRFAVPRPT--GETVMRHQTLPR 301
DB 401 PAAYDAQNDIILYLP-----MGVST-----PDINGGAVTRKRANLPA 436
QY 302 DNM---DOCTEFEMAVTNVDOPSTEGLOSINPNA-----TGERRVILTVCPCKTGTM 353
DB 437 VYVLCMPOLASMHGFTTRYTQMSFMDM-LPS--OPTLADITDEOGKIVPVYVPAKIGNI 494
QY 354 WQPAETNGE-FLMARDT-----NYQMIISIDENGIVTVNED 389
DB 495 FVLNNDIGKVPVAPETVPOGPAKGHLSPOTPSSELTFRPKKILQGRDMWG-ATWEDQ 553

QY 390 ALKEL--DVEYD-----VCPTFLGGRDMSALNPDSGITFIPLNVCYDMM 435
DB 554 LMCVRMFHKLRYEGFPFPSPSEOGTLVFPGDGFMWEGMGSIVN----- 595
QY 436 AVDOEFEMDYNTSNVTKL-----PPGKD----- 460
DB 596 -TDQOFALANMAMPFISKILPRGPNPIEGADGAAGSGSGSVQIMYGVPIVELNPF 654
QY 461 -----MIGRIDAIDISTGRTLWSYRAANYS-PV-----LSTGG 494
DB 655 LSPGLPCLQPSWGFVSALINRNHQLIKKRIQTRPSAPVPLPFKKGVPMLGPTTAG 714
QY 495 GYLFENGCF-DRYFPAALSGEETLMQRLNVAASGA--ISYEDNGQYVAI--AGGSVY 550
DB 715 NIFVAGLIDNLYRAYSVDGSKLLMQARLP--AGGQATPMTEYDQYVYVIMAGHGSI 772
QY 551 GSGINSALAGERVDSITAGNAVVEALP 578
DB 773 G-----TRLGOSLIYKLP 786

RESULT 13
ID 052551 PRELIMINARY; PRT; 644 AA.
AC 052551;
DT 01-NOV-1996 (TREMBLEL 01, Created)
DT 01-NOV-1996 (TREMBLEL 01, Last sequence update)
DT 01-OCT-2000 (TREMBLEL 15, Last annotation update)
DE DEHYDROGENASE.
GN PVA.
OS Pseudomonas sp.
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=306;
RN [1]
RP SEQUENCE FROM N.A.
RA Matsumiya A., Hatemake T., Tsuji M., Takakura K., Takizawa N.,
RA Kiyohara H.,
RT "Molecular Cloning and Nucleotide sequence a Gene that Encoding
RT Poly(Vinyl alcohol) Dehydrogenase from Pseudomonas sp. 113p3."
RL Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL: D83772; BAA12104.1;
DR InterPro: IPR000345;
DR InterPro: IPR002372;
DR InterPro: IPR003088;
DR Pfam: PF00034; Cytochrome_c; 1.
DR Pfam: PF01011; Bacterial_POQ; 6.
DR PROSITE: PS00190; CYTOCHROME_C; 1.
SQ SEQUENCE 644 AA; 69324 MW; 9D817A59DDAA5303 CRC64;

Query Match 9.1%; Score 281.5; DB 2; Length 644;
Best Local Similarity 24.0%; Pred. No. 1e-11;
Matches 144; Conservative 78; Mismatches 218; Indels 161; Gaps 30;

QY 7 LMASAGALLAALAPAFQAQ--TPVTDELLANPPAGEISYSGQDENVRHSPLTOITTEVNG 65
DB 124 IKGNSVEGTPLDPCSSAPFV--DLGA---ANQNGWSTEDNGRFORKALDVADLP 178
QY 66 QIQLVNARGMQPKQVOTPLIHGVMYLANPGDVIOAIDAKTGDLIWEHRRO-----LPNI 121
DB 179 KKLKMA--FGYPSKNGQATVIGDRLFTTSGAVYALNAKGVYWRHRREARTSPVI 237
QY 122 ATINSFGEPTRGALYGVNYPVSNHNLVALDTATG-----QVTFVDVDRG 167
DB 238 AALPE-----ARHKITALFESDFTKAVALDAETGKOLMTVVDDQPALOMTGSITW 289
QY 168 OGEDWYNSGSPYIANGVIVAGSTOYSPFGFVSGHDSATGEELMRNFI--PR---A 221
DB 290 DGIIVYPISSG--TEAPAQIPTECCAFR--GALVA--LDATGKILMKRTTTEQRPPEKL 345
QY 222 GEGDETWGNDYFARMWTGAMGOITYDPVTNLVHYGSTAVGPASSETORGTPGGLYGTNT 281

```

Db 346 NKAGROMGPGSGAIWTP-----TVDEARLIYGT-----SNSTIYVP---YDNSDS 391
QY 282 RFAVRPTGELIWMRHOTLPDNDDECTFEEMVTNVDPSTEGHQS-INPNAATGER 340
Db 392 VMAIDADTGAIVTVOGLADNVIIDGCM-----QKGKEHIANPNLPGPFSIGRA 441
QY 341 RVL-----TGV-PCKGTGMQFPAETGEFLMA--RDTNTQNMIESIDENGIVTVNEDAI 391
Db 442 PYLEDGGRKVSPCR-----AEIGHDLKARFRPTFARREWE-----478
QY 392 LKEIDVEXDPTFLGGGRDMPALNDPSGIYFPLNNVCYDMAVDOEFTSMVNTSN 451
Db 479 -RQLSL-----GSLGMEGETTA---DDQKVYAGVSDIA--SOAKDRG-----516
QY 452 VTKLPPGKMDIGRIDAIDISTGRTLMSVERAA-----ANSPVLSTGGVLFN 499
Db 517 -----KPG-----LMALDIRTGEVAMNPLNAPDKCRNMNMWCHGARSQALSVIPGIFA 566
QY 500 GGTGRTFRALSQETGELTMO-----TRLATYASGCAISYEVDGMQYVLAAGGVSYSGLN 555
Db 567 GSYDGHFRAEDPTATGKIIMVDGTGKAVTTLG-----AKAFGGVMDGAGR 613
QY 556 S 556
Db 614 S 614

RESULT 14
P77931 ID P77931 PRELIMINARY; PRT; 639 AA.
AC P77931;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)
DE POO-DEPENDENT POLYVINYL ALCOHOL DEHYDROGENASE PRECURSOR
DE (EC 1.1.99.23).
CN PVA.
OS Pseudomonas sp.
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=306;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VM15C;
MEDLINE=96376165; PubMed=8782398;
RA Shimo M., Tamogami T., Nishi K., Harayama S.;
RT "The gene pvaB encodes oxidized polyvinyl alcohol hydrolase of
RT Pseudomonas sp. strain VM15C and forms an operon with the polyvinyl
RT alcohol dehydrogenase gene pvaA.";
RL Microbiology 146:649-657(2000).
DR EMBL: D50670; BAA09321.1; -.
DR EMBL: AB008494; BAA94193.1; -.
DR InterPro: IPR000345; -.
DR InterPro: IPR002372; -.
DR InterPro: IPR003088; -.
DR Pfam: PF00034; cytochrome_c_1.
DR Pfam: PF01011; Bacterial_POO; 5.
DR PROSITE: PS00190; CYTOCHROME_C_1.
KW Signal; Oxidoreductase; POO.
FT SIGNAL 1 33 POTENTIAL.
SQ SEQUENCE 639 AA; 68049 MW; 406E9EF873963B8A CRC64;

```

Query Match 9.08; Score 276.5; DB 2; Length 639;
Best Local Similarity 22.48; Pired. No. 2.3e-11;

```

Matches 143; Conservative 70; Mismatches 241; Indels 183; Gaps 28;
QY 7 LMASAGALLAAPAAQVPTDELLANPAGEWISYQONENYHSPLOTITENNGO 66
Db 117 IWGPSPASMPIDSPKCKKGIIPID--LSTP--DQWNGWAGITNARFQNPGLTADVR 172
QY 67 LQLVARGMQPGKQVQVTPLIHDSVMTLANPGVIOAIDAKTGDLIWEHRQLPNIATLS 126
Db 173 LKYMAMFN-YPGSKNQAVTVGDRLEFVTSMSGAVYALNAKTGCYVRHDAATASSVAV 231
QY 127 FGSPTGMLLYGNVFEVMDNLVALDTATG---QVFEDVDRG-----QGEEM 172
Db 232 VOLPAGAPAOYA--IFSPWTKAVALDAOTGOLKTTIIDQPGVOMTGSPTYHGGKLP 289
QY 173 VSNSSGPPIYANGVYAGSTQYSPFCFVSGHDSATGEELMENVYIP-----RAGEED 226
Db 290 VPISSG-----NEAFNNDQWECKFRGALVALDALSGKVLKMTYTTQKRPAPFLNKLK 345
QY 227 ETWGNDEYARMMTGA--WQIITYDPVTNLV-----HYSTAVGPASEFTQRTP 272
Db 346 OMWG-----PAGSISWAPTIIDPKRGLVYVATSNSTYEVHSGSDAV-----387
QY 273 GGTLYGTNTRFAVRPTGELIWMRHOTLPDNDDECTFEEMVTNVDPSTEGHQSIN 332
Db 388 -----MAELETGKVRMINOYTKDNTYIGCP--RANCP-----EKVG 424
QY 333 PNAATGERVLL-----TGVPCKTGTMOFDAET-GEFLMARDNTQNMIESIDENG 382
Db 425 PDEALGNSPILHTLDGQRIYVGGKSGAVYAMPDNDSELIWMR-----469
QY 383 IYTVNEDAILKELDEYDVCPFLGGDMPSAALNDPSGIYFPLNNVCYDMAVDOEFT 442
Db 470 -----RVSPGSEL--GGVER-----GMAADAE--489
QY 443 SMDVY-NTSNVTKLPKPGKMIGRIDAIDISTGRTLS--VEEAANY-----SPVL 490
Db 490 --NVYGISDVITTRKGGK--GYVALRIDADVWAPRPTPCRMNIFCHPAVSQAV 544
QY 491 STGGGVLPNGGTDYRFRALSQETGELTMOETRLA-----TVASGOAISYEVDGMQYVLA 545
Db 545 TAMPGVYFAGSMDGHRFRASTSDGKVLMEFNFAAPYKTVAAKQADGGMDSAG-PTIAG 603
QY 546 GGVSYSGSLNSALAGERVDSTA-----IGNAVYVAL 577
Db 604 GMVYVHSG-----YAGRSTQAGDLRGEGNVLAIRSV 636

RESULT 15
ID 030326 PRELIMINARY; PRT; 470 AA.
AC 030326;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, last annotation update)
DE GLUCOSE DEHYDROGENASE-A [PYRROLOQUINOLINE-QUINONE] (EC 1.1.99.17)
DE (QUINOPROTEIN GLUCOSE DH) (GDH-A) (FRAGMENT).
GN GDHA.
OS Acetobacter europaeus.
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
OC Glucanacetobacter.
OX NCBI_TaxID=33995;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 6160;
MEDLINE=97385083; PubMed=9238099;
RA Thuner C.A., Vela C., Theony-Meyer L., Melle L., Teuber M.;
RT "Biochemical and genetic characterization of the acetaldehyde
RT dehydrogenase complex from Acetobacter europaeus.";
RL Arch. Microbiol. 168:81-91(1997).
CC -I- CATALYTIC ACTIVITY: D-GLUCOSE + ACCEPTOR = D-GLUCONO-1,5-LACTONE +
CC + REDUCED ACCEPTOR
CC -I- CORFACTOR: POO (BY SIMILARITY).
CC -I- SUBUNIT: MONOMER (BY SIMILARITY).

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 9, 2001, 10:10:01 ; Search time 72.18 Seconds
(without alignments)
4563.610 Million cell updates/sec

Title: US-08-934-506A-1

Perfect score: 1740
Sequence: 1 ATGAACCGACTGCTGCT.....TCTTGGCCCTGCCCAATTA 1740

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 segs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: Issued_Patents_NA:*

1: /cgn2_6/prodata/1/lna/5A.COMB.seq:*

2: /cgn2_6/prodata/1/lna/5B.COMB.seq:*

3: /cgn2_6/prodata/1/lna/5A.COMB.seq:*

4: /cgn2_6/prodata/1/lna/5B.COMB.seq:*

5: /cgn2_6/prodata/1/lna/5A.COMB.seq:*

6: /cgn2_6/prodata/1/lna/5B.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	92.2	5.3	2214	1	US-07-985-458-1	Sequence 1, Appl
2	76.2	4.4	2163	4	US-09-286-284-22	Sequence 22, Appl
3	76.2	4.4	2265	4	US-09-286-284-1	Sequence 1, Appl
4	76.2	4.4	4830	4	US-09-286-284-7	Sequence 7, Appl
5	42.2	2.4	2745	1	US-08-363-255-1	Sequence 1, Appl
6	42.2	2.4	2745	1	US-08-363-255-13	Sequence 13, Appl
7	41.4	2.4	655	1	US-08-471-119A-5	Sequence 5, Appl
8	40.2	2.3	44377	2	US-08-804-227C-7	Sequence 7, Appl
9	40.2	2.3	44377	2	US-08-804-198-1	Sequence 1, Appl
10	39.8	2.3	1164	1	US-07-640-476-6	Sequence 6, Appl
11	39.4	2.3	3748	2	US-08-666-417-2	Sequence 2, Appl
12	39.4	2.3	36519	3	US-08-923-137-2	Sequence 98, Appl
13	39.4	2.3	5515	4	US-09-398-193-8	Sequence 3, Appl
14	38.8	2.2	20235	1	US-07-642-734C-3	Sequence 3, Appl
15	38.8	2.2	20235	3	US-08-439-009A-3	Sequence 18, Appl
16	38.4	2.2	2004	1	US-08-471-033-18	Sequence 18, Appl
17	38.4	2.2	2004	2	US-08-471-044-18	Sequence 18, Appl
18	38.4	2.2	2004	2	US-08-463-483A-18	Sequence 18, Appl
19	38.4	2.2	2004	2	US-08-471-046A-18	Sequence 18, Appl
20	38.4	2.2	2004	2	US-08-470-566B-18	Sequence 18, Appl
21	38.4	2.2	2004	2	US-08-469-334-18	Sequence 18, Appl
22	38.4	2.2	2004	3	US-09-300-529-18	Sequence 18, Appl
23	38.4	2.2	2576	1	US-08-471-033-35	Sequence 35, Appl
24	38.4	2.2	2576	2	US-08-471-044-35	Sequence 35, Appl
25	38.4	2.2	2576	2	US-08-463-483A-35	Sequence 35, Appl
26	38.4	2.2	2576	2	US-08-471-046A-35	Sequence 35, Appl
27	38.4	2.2	2576	2	US-08-470-566B-35	Sequence 35, Appl

28	38.4	2.2	2576	2	US-08-469-334-35	Sequence 35, Appl
29	38.4	2.2	2576	3	US-09-300-529-35	Sequence 35, Appl
30	38.4	2.2	2655	1	US-08-471-033-17	Sequence 17, Appl
31	38.4	2.2	2655	1	US-08-471-033-26	Sequence 26, Appl
32	38.4	2.2	2655	2	US-08-471-044-17	Sequence 17, Appl
33	38.4	2.2	2655	2	US-08-471-044-26	Sequence 26, Appl
34	38.4	2.2	2655	2	US-08-463-483A-17	Sequence 17, Appl
35	38.4	2.2	2655	2	US-08-463-483A-26	Sequence 26, Appl
36	38.4	2.2	2655	2	US-08-471-046A-17	Sequence 17, Appl
37	38.4	2.2	2655	2	US-08-471-046A-26	Sequence 26, Appl
38	38.4	2.2	2655	2	US-08-470-566B-17	Sequence 17, Appl
39	38.4	2.2	2655	2	US-08-470-566B-26	Sequence 26, Appl
40	38.4	2.2	2655	2	US-08-469-334-17	Sequence 17, Appl
41	38.4	2.2	2655	2	US-08-469-334-26	Sequence 26, Appl
42	38.4	2.2	2655	3	US-09-300-529-17	Sequence 17, Appl
43	38.4	2.2	2655	3	US-09-300-529-26	Sequence 26, Appl
44	38.4	2.2	3407	1	US-08-253-155A-7	Sequence 7, Appl
45	38.4	2.2	4031	1	US-08-471-033-49	Sequence 49, Appl

ALIGNMENTS

RESULT 1
US-07-985-458-1
Sequence 1, Application US/07985458
Patent No. 5344777
GENERAL INFORMATION:
APPLICANT: Takemura, Toshimi;
APPLICANT: Takemura, Hiroshi;
APPLICANT: Takemura, Kenji;
APPLICANT: Fukaya, Masahiro;
APPLICANT: Okumura, Hajime and
APPLICANT: Kawamura, Yoshiya
TITLE OF INVENTION: Structural Gene of Membrane-bound
TITLE OF INVENTION: Alcohol Dehydrogenase Complex, Plasmid
TITLE OF INVENTION: Containing The Same And Transformed Acetic Acid
TITLE OF INVENTION: Bacteria
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: Frischauf, Holztz, Goodman & Woodward, P.C.
STREET: 600 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10016-2088
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 0.72 mb
COMPUTER: IBM PC compatible (NEC PC-9801 ES)
OPERATING SYSTEM: MS DOS
SOFTWARE: ASCII Form
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/985,458
FILING DATE: 19921203
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/658,221
FILING DATE: 20-FEB-1991
APPLICATION NUMBER: 73440/1990
FILING DATE: 26-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Goodman, Herbert
REGISTRATION NUMBER: 17081
REFERENCE/DOCKET NUMBER: 910134/HG
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)972-1400
TELEFAX: (212)370-1622
TELEX: 236268
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2214 base pairs
TYPE: nucleic acid
STRANDEDNESS: double

Db 190 agcccgctgagatcagatcaccagaagacatgagcaatctgaagctgagctgcaactac 249
QY 220 GGCAATGACAGCGGGCAAGATC---AAGTACAGCCCTGATTCATGACGGCGCATGATAT 276
Db 250 gatctgatalccaccacgctggtcaggaagtagcagcctgactgctgtagcgtcatgtac 309
QY 277 CTGGCAAAACCGGGCGAGCATGATCAGGACGCAAAACCTGGCGATTCGATCTG 336
Db 310 gccacacaacatgagcaagaatgaagcctctgtagcagctgaagcagcagctgctgag 369
QY 337 GAACACCGCGCCCAACTGCC---GAACATGCCACAGCTTGAACGCTTGGCGACCGAC 393
Db 370 tcttaagatccaaaggttccaggaacatcgcgacccgctgctgctgagtagcgtcaac 429
QY 394 CGCGGCAATGGCGCTGTACGGCACCACGTTTACTTGTGTTGCGGACAACACCTGCTG 453
Db 430 cgtgtgtagcagctactggaacggaagctatcttcgcaaccttcgagcgctgcctgatt 489
QY 454 GCCCTGCACACCGCAACTGGCGCAAGTGCCTGACGTCGA 494
Db 490 gccctgtagtccaaagcagcgagctgctgtagcgtcta 530

RESULT 4

US-09-296-284-7
; Sequence 7, Application US/09296284A
; Patent No. 6204040
; GENERAL INFORMATION:
; APPLICANT: Choi, Eun-Sung
; APPLICANT: Rhee, Sang-Ki
; APPLICANT: Lee, Eun-Hae
; TITLE OF INVENTION: Glucobacter Suboxydans Sorbitol Dehydrogenase, Genes
; FILE REFERENCE: 1533.0870000
; CURRENT APPLICATION NUMBER: US/09/296,284A
; CURRENT FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 7
; LENGTH: 4830
; TYPE: DNA
; ORGANISM: Glucobacter suboxydans
US-09-296-284-7

Query Match 4.4%; Score 76.2; DB 4; Length 4830;
Best Local Similarity 52.9%; Pred. No. 1.2e-09;
Matches 212; Conservative 0; Mismatches 183; Indels 6; Gaps 2;

QY 100 GCGAACCAGCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 159
Db 794 gccgaccagatccggtgtagctgtagctgtagctgtagctgtagctgtagctgtagct 853
QY 160 TCGGCCCCGACGATCAGATCAGATCAGATCAGATCAGATCAGATCAGATCAGATC 219
Db 854 agcccgctgtagatcaccacgaagacatgagcaatctgaagctgtagctgtagctgtag 913
QY 220 GGCATGACAGCGGGCAAGATC---AAGTACAGCCCTGATTCATGACGGCGCATGAT 276
Db 914 gatctgtagtaccacgctgtagcaggaagtagcagcgtgtagttagttagttagttagt 973
QY 277 CTGGCAAAACCGGGCGAGCATGATCAGGCGCATGACGCCAAACCTGGCGATTCGATCTG 336
Db 974 gccacacaacatgagcaagaatgaagcctctgtagcagctgtagcagcagcagctgtag 1033
QY 337 GAACACCGCGCCCAACTGCC---GAACATGCCACAGCTTGAACGCTTGGCGACCGAC 393
Db 1034 tcttaagatccaaaggttccaggaacatcgcgacccgctgctgtagtagcgtcaac 1093
QY 394 CGCGGCAATGGCGCTGTACGGCACCACGTTTACTTGTGTTGCGGACAACACCTGCTG 453
Db 1094 cgtgtgtagcagctactggaacggaagctatcttcgcaaccttcgagcgctgcctgatt 1153

QY 454 GCCCTGCACACCGCAACTGGCGCAAGTGCCTGACGTCGA 494
Db 1154 gccctgtagtccaaagcagcgagctgctgtagcgtcta 1194

RESULT 5

US-08-363-255-1
; Sequence 1, Application US/08363255
; Patent No. 5783386
; GENERAL INFORMATION:
; APPLICANT: JACOBS, JR., WILLIAM R.
; APPLICANT: BLOOM, BARRY R.
; APPLICANT: COLLINS, DESMOND M.
; APPLICANT: de LISLE, GEOFFREY W.
; APPLICANT: PASCOBELLA, LISA
; APPLICANT: KAMAKAMI, RIKU P.
; TITLE OF INVENTION: MYCOBACTERIA VIRULENCE FACTORS AND A
; TITLE OF INVENTION: NOVEL METHOD FOR THEIR IDENTIFICATION
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/363,255
; FILING DATE: 23-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MONROY, GLADYS H.
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 25237-20002.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2745 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 835..2424
US-08-363-255-1

Query Match 2.4%; Score 42.2; DB 1; Length 2745;
Best Local Similarity 45.4%; Pred. No. 0.18;
Matches 152; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

QY 985 CAGTGCATCAACCCGACAGCGCCCAACTGGCGAGCGTCCGTGCTGACGGCGTCCGTGC 1044
Db 1006 CGGTCCGCTCAACCCCGCTCGGACCCGACGACACTAGACACCATCCCGAAGG 1065
QY 1045 AAACCCGACCATGTGGAGTTCGACGCGCGAAGCCGCGAATTCCTGTGGCCCGTGTAT 1104
Db 1066 AAGACCCCGCGCGCGCCCAATTCGCCCGCGGAGAGGACCGTGGCCCGCGCACGCG 1125
QY 1105 ACCAATCAACAGACATGATGATCAATCAAGCAAGAAAGGCGCATGTCACCGTAACGA 1164
Db 1126 ACCAAGCCACGGGCGCCCAAGGATGCGCAGCAGAAAGCGGACAGAGATCCCGAGAGCG 1185
QY 1165 GATGCAATCCGAAGAACTGATGATTTGAATATGACGCTGCGCCGACCTTTGTTGGCGGC 1224

NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: PAUL R. CANTRELL, 1138
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,198
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CANTRELL, PAUL R.
REGISTRATION NUMBER: 36,470
REFERENCE/DOCKET NUMBER: P9113
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3885
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4437 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 350..14002
FEATURE:
NAME/KEY: CDS
LOCATION: 14046..20036
FEATURE:
NAME/KEY: CDS
LOCATION: 20110..31284
FEATURE:
NAME/KEY: CDS
LOCATION: 31329..36071
FEATURE:
NAME/KEY: CDS
LOCATION: 36155..41830
US-08-804-198-1

Query Match 2.3%; Score 40.2; DB 2; Length 44377;
Best Local Similarity 41.8%; Pred. No. 1.2;
Matches 246; Conservative 0; Mismatches 343; Indels 0; Gaps 0;

QY 1134 CGACGAAAGGCGATGTACCGGTGACGACGACGATGCTGGAAGAACTGGATGTTGA 1193
DB 637 CGCGGCGATCGTCCCGGCTCTGCGCGCGACGAGGGGCGCGCTTCGTGCGGGCCAT 696
QY 1194 ATATGACGTCGCGCGATCTTGTGGGGCGCGACGTGCGCGCGCGACGTGAACCC 1253
DB 697 GCACGACGACTGACGACCGCTGCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCG 756
QY 1254 GCAGACGCGGCGATCTACTTCTATCCGCTGAACAGCTGTGCTATGACATGATGGCGTGA 1313
DB 757 CACCGGCTTCCAGCGCGCGCGCTGCGCCAAAGCGGCTCTCTAGCTGCTGGGAGCGCGGG 816
QY 1314 TCAGGATTCACCTCGATGAGAGCTCTATTAACACGACGACGACGACGACGACGACG 1373
DB 817 CCCCACTTCGCGGCGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 876
QY 1374 CAAGGATATGATGCGTGTATTGACGCGATGACATGACGACGACGACGACGACGACG 1433
DB 877 CGAGAGCTTGGCG 936
QY 1434 GGTGACGCTGCTGGCGCGACGACGACGACGCGCTTGTGACCGCGCGCGCGCGCGCT 1493

DB 937 CGCCGACGAGGAAAGCGCGCCCATGGAACGCGCTGCGCGCGCTGCACCGCGCGCGCTG 996
QY 1494 CACGCGTGTACGATGCTGTTCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1553
DB 997 CCACACCTTCGACGCGCGCTGCGCAACGCGCTATGCTCGGAGGCGCGCGCGCGCGCT 1056
QY 1554 GCAGACCGCGCTTGCACGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1613
DB 1057 CCGTGAAGCCCTCG 1116
QY 1614 GCAATATGCG 1673
DB 1117 CGTGCCTCGCGCAACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1176
QY 1674 GGTGCGGCGAGGCG 1722
DB 1177 GGAGCGGTGCTCCG 1225

RESULT 10
US-07-640-476-6
Sequence 6, Application US/07640476
Patent No. 5376536
GENERAL INFORMATION:
APPLICANT: QUAX, WILHELMUS
APPLICANT: LUITEN, RUDOLF G.M.
APPLICANT: SCHUIRHUIZEN, PAUL W.
APPLICANT: MRABET, NADIR
TITLE OF INVENTION: NOVEL GLUCOSE ISOMERASE ENZYMES AND
TITLE OF INVENTION: THEIR USE
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 545 Middlefield Road, Suite 200
CITY: Menlo Park
STATE: CA
COUNTRY: USA
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/640,476
FILING DATE: 19910110
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kate H. Murashige
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 24615-20009.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 327-7250
TELEFAX: (415) 327-2951
TELEX: 706141
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1164 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORGANISM: Streptomyces murinus
STRAIN: DSM 40091
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1164
IDENTIFICATION METHOD: experimental

Query Match	2.2%	Score 39	DB 4	Length 5515
Best Local Similarity	44.0%	Pred. No. 1,3'		
Matches 165	Conservative	0	Mismatches 210	Indels 0
Gaps				0
QY	982	CTGCAGTCATCTCAACCGGACGCGCAATCTGGCGACGTCCGCTCTACACCGGGCTTCG	1041	
Db	3632	ctgaggaacatcatccctaccacgcytfgtctgacgcctggaagtygttccgaacttaactc	3691	

BIOPHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Saccharopolyspora erythraea
STRAIN: NRRL 238
FEATURE:


```

1 LOCATION: 97..1482 /function= "approximate span of
2 OTHER INFORMATION: beta-ketoacyl ACP synthase of module 3"
3 OTHER INFORMATION: beta-ketoacyl ACP synthase of module 3"
4 FEATURE:
5 NAME/KEY: misc_feature
6 LOCATION: 1693..2670 /function= "approximate span of
7 OTHER INFORMATION: acyltransferase domain module 3"
8 OTHER INFORMATION: acyltransferase domain module 3"
9 FEATURE:
10 NAME/KEY: misc_feature
11 LOCATION: 3405..3921 /function= "approximate span of
12 OTHER INFORMATION: beta-ketoreductase domain of module 3"
13 OTHER INFORMATION: beta-ketoreductase domain of module 3"
14 FEATURE:
15 NAME/KEY: misc_feature
16 LOCATION: 4171..4428 /function= "approximate span of
17 OTHER INFORMATION: acyl carrier domain of module 3"
18 OTHER INFORMATION: acyl carrier domain of module 3"
19 FEATURE:
20 NAME/KEY: misc_feature
21 LOCATION: 4471..10722 /function= "approximate span of
22 OTHER INFORMATION: module 4"
23 OTHER INFORMATION: module 4"
24 FEATURE:
25 NAME/KEY: misc_feature
26 LOCATION: 4471..5847 /function= "approximate span of
27 OTHER INFORMATION: beta-ketoacylACPSynthase domain of module
28 OTHER INFORMATION: beta-ketoacylACPSynthase domain of module
29 FEATURE:
30 NAME/KEY: misc_feature
31 LOCATION: 6054..7026 /function= "approximate span of
32 OTHER INFORMATION: acyltransferase domain of module 4"
33 OTHER INFORMATION: acyltransferase domain of module 4"
34 FEATURE:
35 NAME/KEY: misc_feature
36 LOCATION: 7165..9216 /function= "approximate span of
37 OTHER INFORMATION: dehydratase and enoylreductase domains m"
38 OTHER INFORMATION: dehydratase and enoylreductase domains m"
39 FEATURE:
40 NAME/KEY: misc_feature
41 LOCATION: 9433..9984 /function= "approximate span
42 OTHER INFORMATION: beta-ketoreductase of module 4"
43 OTHER INFORMATION: beta-ketoreductase of module 4"
44 FEATURE:
45 NAME/KEY: misc_feature
46 LOCATION: 10225..10483 /function= "approximate span of
47 OTHER INFORMATION: acyl carrier domain of module 4"
48 OTHER INFORMATION: acyl carrier domain of module 4"
49 FEATURE:
50 NAME/KEY: CDS
51 LOCATION: 10723..20235 /codon_start= 10723
52 OTHER INFORMATION: /codon_start= 10723
53 OTHER INFORMATION: /function= "gene -erya"
54 OTHER INFORMATION: /product= "orf3 encoding modules 5 & 6
55 OTHER INFORMATION: 6-deoxyerythronolide B formatto"
56 FEATURE:
57 NAME/KEY: misc_feature
58 LOCATION: 10723..15165 /function= "approximate span of
59 OTHER INFORMATION: module 5"
60 OTHER INFORMATION: module 5"
61 FEATURE:
62 NAME/KEY: misc_feature
63 LOCATION: 10831..12174 /function= "approximate span of
64 OTHER INFORMATION: beta-ketoacylACPSynthase domain of modul"
65 OTHER INFORMATION: beta-ketoacylACPSynthase domain of modul"
66 FEATURE:
67 NAME/KEY: misc_feature
68 LOCATION: 12379..13350 /function= "approximate span of
69 OTHER INFORMATION: acyltransferase domain of module 5"
70 OTHER INFORMATION: acyltransferase domain of module 5"
71 FEATURE:
72 NAME/KEY: misc_feature
73 LOCATION: 14062..14610 /function= "approximate span of
74 OTHER INFORMATION: module 5"
75 OTHER INFORMATION: module 5"

```


Search completed: August 9, 2001, 22:38:11
Job time: 44890 sec

```

OTHER INFORMATION: /function="approximate span of
OTHER INFORMATION: beta-ketoreductase of module 5"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 14857..15114
OTHER INFORMATION: /function="approximate span of
OTHER INFORMATION: acyl carrier domain of module 5"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 15166..20235
OTHER INFORMATION: /function="approximate span of
OTHER INFORMATION: module 6"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 15172..16569
OTHER INFORMATION: /function="approximate span of
OTHER INFORMATION: beta-ketocacylCPSynthase domain of modul"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 16768..17721
OTHER INFORMATION: /function="approximate span of
OTHER INFORMATION: acyltransferase domain of module 6"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 18379..18921
OTHER INFORMATION: /function="approximate span of
OTHER INFORMATION: beta-ketoreductase domain of module 6"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 19149..19398
OTHER INFORMATION: /function="approximate span of
OTHER INFORMATION: acyl carrier domain of module 6"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 19492..20235
OTHER INFORMATION: /function="approximate span of
OTHER INFORMATION: thioesterase domain of module 6"
US-08-439-009A-3

```

```

Query Match          2.2%; Score 38.8; DB 3; Length 20235;
Best Local Similarity 45.9%; Pred. No. 2.1; Mismatches 197; Indels 1; Gaps 1;
Matches 168; Conservative 0;
QY 335 GGGACACCCCGCCCAACTCCGACATCCGACGCTGACAGCTTTGGCG-AGCCGACC 393
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6304 GTACACCCCGCCCGCGCTGTCGGCCACTCGCAGGGGAGATCGCCGCGCACGTGGCG 6363
QY 394 CGCGGCGATGGCGCTGTACGGCCCAACGTTTACTTTGTTGTTGGACACCACTGGTC 453
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6364 GCGCGCGTGTGCTGAGACCGCCGCAAGGTGTGCGCCCTGCGCAGCCAGTGTGGCG 6423
QY 454 GCCCTGACACCGCACTGGCCCAAGTACGTGTGACGTGACCGCGGCCCAAGGCAAGAC 513
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6424 GAGTCTGACGACGAGGCGGCGATGTGTGCGCGCGCTCCCGCAGCAGCTGGAGACC 6483
QY 514 ATGGTTTCGAATCTGTCGGGCGCCGATCGTGTGCAACGAGCGGTGATCGTGGCGTTGAC 573
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6484 GTGCTCGCGCGCTGGAGCGCCGTGTGCGCGTGGCGCGCTGACAGCGCTGTGACACGAC 6543
QY 574 TGCCTACTCGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 633
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6544 GTGCTTCCGCGCGCGCGGAGCTGAGAGTCTTCTGCGCAGGCGCGGCGCGGAG 6603
QY 634 CTGTGGGCGCACTTACTTATCCCGCGCTGGGAGAGAGGTATGAGACTTGGGCAAC 693
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6604 ATGAAGCCGCGCGGATCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6663
QY 694 GATTAC 699
    || ||
Db 6664 GAGGAC 6669

```



GenCore version 4.5
Copyright (c) 1993 - 2000 Compen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 9, 2001, 09:52:11 ; Search time 17754.2 seconds
(without alignments)
926.427 Million cell updates/sec

Title: US-08-934-506A-1

Perfect score: 1740

Sequence: 1 ATGAACCGACTGCTGCT.....TCCTGCCCTGCCGCAATAA 1740

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 segs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: gb_estba:*
34: gb_estfun:*
35: gb_esthum1:*
36: gb_esthum2:*
37: gb_esthum3:*
38: gb_esthum4:*
39: gb_esthum5:*
40: gb_esthum6:*
41: gb_esthum7:*
42: gb_esthum8:*
43: gb_esthum9:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estp11:*
73: em_estp12:*
74: em_estp13:*
75: em_estp14:*
76: em_estp15:*
77: em_estp16:*
78: em_estp17:*
79: em_estp18:*
80: em_estp19:*
81: em_estp110:*
82: em_estp110:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_estc25:*
103: gb_estc26:*
104: gb_estc27:*
105: gb_estc28:*
106: gb_estc29:*
107: gb_estc30:*
108: gb_estc31:*
109: gb_estc32:*
110: gb_estc41:*
111: gb_estc42:*
112: gb_estc43:*
113: gb_estc44:*
114: gb_estc45:*
115: gb_estc46:*
116: gb_estc47:*

117: gb_estf48:*
118: gb_estf49:*
119: gb_estf50:*
120: gb_estf51:*
121: gb_estf52:*
122: gb_estf53:*
123: gb_estf54:*
124: gb_estf55:*
125: gb_estf56:*
126: gb_estf57:*
127: gb_estf58:*
128: gb_estf59:*
129: gb_estf60:*
130: gb_estf61:*
131: gb_estf62:*
132: gb_estf63:*
133: gb_estf64:*
134: gb_estf65:*
135: gb_estf66:*
136: gb_estf67:*
137: gb_estf68:*
138: gb_estf69:*
139: gb_estf70:*
140: gb_estf71:*
141: gb_estf72:*
142: gb_estf73:*
143: gb_estf74:*
144: gb_estf75:*
145: gb_estf76:*
146: gb_estf77:*
147: gb_estf78:*
148: gb_estf79:*
149: gb_estf80:*
150: gb_estf81:*
151: gb_estf82:*
152: gb_estf83:*
153: gb_estf84:*
154: gb_estf85:*
155: gb_estf86:*
156: gb_estf87:*
157: gb_estf88:*
158: gb_estf89:*
159: gb_estf90:*
160: gb_estf91:*
161: gb_estf92:*
162: gb_estf93:*
163: gb_estf94:*
164: gb_estf95:*
165: gb_estf96:*
166: gb_estf97:*
167: gb_estf98:*
168: gb_estf99:*
169: gb_estf00:*
170: gb_estf01:*
171: gb_estf02:*
172: gb_estf03:*
173: gb_estf04:*
174: gb_estf05:*
175: gb_estf06:*
176: gb_estf07:*
177: gb_estf08:*
178: gb_estf09:*
179: gb_estf10:*
180: gb_estf11:*
181: gb_estf12:*
182: gb_estf13:*
183: gb_estf14:*
184: gb_estf15:*
185: gb_estf16:*
186: gb_estf17:*
187: gb_estf18:*
188: gb_estf19:*
189: gb_estf20:*

190: gb_estl10:*
191: gb_estl11:*
192: gb_estl12:*
193: gb_estl13:*
194: gb_estl14:*
195: gb_estl15:*
196: gb_estl16:*
197: gb_estl17:*
198: gb_estl18:*
199: gb_estl19:*
200: gb_estl20:*
201: gb_estl21:*
202: gb_estl22:*
203: gb_estl23:*
204: gb_estl24:*
205: gb_estl25:*
206: gb_estl26:*
207: gb_estl27:*
208: gb_estl28:*
209: gb_estl29:*
210: gb_estl30:*
211: gb_estl31:*
212: gb_estl32:*
213: gb_estl33:*
214: gb_estl34:*
215: gb_estl35:*
216: gb_estl36:*
217: gb_estl37:*
218: gb_estl38:*
219: gb_estl39:*
220: gb_estl40:*
221: gb_estl41:*
222: gb_estl42:*
223: gb_estl43:*
224: gb_estl44:*
225: gb_estl45:*
226: gb_estl46:*
227: gb_estl47:*
228: gb_estl48:*
229: gb_estl49:*
230: gb_estl50:*
231: gb_estl51:*
232: gb_estl52:*
233: gb_estl53:*
234: gb_estl54:*
235: gb_estl55:*
236: gb_estl56:*
237: gb_estl57:*
238: gb_estl58:*
239: gb_estl59:*
240: gb_estl60:*
241: gb_estl61:*
242: gb_estl62:*
243: gb_estl63:*
244: gb_estl64:*
245: gb_estl65:*
246: gb_estl66:*
247: gb_estl67:*
248: gb_estl68:*
249: gb_estl69:*
250: gb_estl70:*
251: gb_estl71:*
252: gb_estl72:*
253: gb_estl73:*
254: gb_estl74:*
255: gb_estl75:*
256: gb_estl76:*
257: gb_estl77:*
258: gb_estl78:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	49.2	2.8	298	246	A2577200	A2577200 04g03 Sh
2	49.2	2.8	925	219	CNS0091P	AL053013 Drosophi
3	46	2.6	643	171	BG937817	BF337817 fme8d03
4	45.6	2.6	667	175	BG307693	BG307693 fme1903
5	45.4	2.6	925	219	CNS0091P	AL053013 Drosophi
6	43.6	2.5	615	121	AM827060	AM827060 fks57h09
7	43.6	2.5	662	175	BG308084	BG308084 fms6f02
8	43.6	2.5	740	137	BE602566	BE602566 HVSME012
9	43.4	2.5	681	220	CNS02FE9P	AL194922 Tetracodo
10	43	2.5	308	164	BE211954	BE211954 894024H00
11	43	2.5	759	146	BF259495	BF259495 HVSMEf001
12	42.8	2.5	880	167	BE456143	BE456143 HVSMEg000
13	42.2	2.4	767	147	BF305223	BF305223 601892676
14	42.2	2.4	553	166	BE364747	BE364747 p11_15 A
15	41.8	2.4	469	155	BG557222	BG557222 EMI_44 B
16	41.8	2.4	487	167	BE429911	BE429911 FAS005 A
17	41.8	2.4	641	77	BE214833	Be214833 HV_CEB000
18	41.8	2.4	1101	219	CNS0175X	AL108460 Drosophi
19	41.6	2.4	586	139	BE758807	BE758807 au_3051 A
20	41.4	2.4	839	219	CNS0004XB	AL054280 Drosophi
21	41.4	2.4	977	219	CNS0003X7	AL076650 Drosophi
22	41	2.4	648	236	AC960102	AC960102 LEREPE38T
23	40.8	2.3	394	155	BG556077	BG556077 EMI_66 D
24	40.8	2.3	476	166	BE364011	BE364011 p11_11 G
25	40.8	2.3	485	164	BE215228	BE215228 HV_CEB000
26	40.8	2.3	499	155	BG557431	BG557431 EMI_44 A
27	40.8	2.3	513	137	BE559448	BE559448 HV_CEB000
28	40.8	2.3	1101	219	CNS0175X	AL108460 Drosophi
29	40.4	2.3	530	136	BE490518	BE490518 HV_CEB000
30	40.4	2.3	456	136	BE497212	BE497212 WHE07362 E
31	40.4	2.3	680	136	BE470555	BE470555 WHE07212
32	40.4	2.3	714	164	BE216552	BE216552 WHE0261_1
33	40.4	2.3	807	167	BE418273	BE418273 HV_CEB000
34	40.4	2.3	878	167	BE417933	BE417933 SCT023_6C
35	40.4	2.3	472	164	BE215776	BE215776 HV_CEB000
36	40.2	2.3	513	164	BE215786	BE215786 HV_CEB000
37	40.2	2.3	517	136	BE500032	BE500032 WHE0977_L
38	40.2	2.3	521	137	BE606263	BE606263 WHE0905_H
39	40.2	2.3	524	167	BE425563	BE425563 WHE0317_G
40	40.2	2.3	528	136	BE499986	BE499986 WHE0977_H
41	40.2	2.3	539	146	BF265039	BF265039 HV_CEA001
42	40.2	2.3	522	136	BE498008	BE498008 WHE0968_L
43	40.2	2.3	594	146	BF263305	BF263305 HV_CEA000
44	40.2	2.3	756	146	BF265359	BF265359 HV_CEA001
45	40.2	2.3	759	146	BF265401	BF265401 HV_CEA001

ALIGNMENTS

	Laboratoire de Biologie Moléculaire des Plantes Supérieures
	University of Geneva
	1 Chemin de l'Imperatrice, Chambesey/Geneva 1292, Switzerland
	Tel : +44(0)1603450000
	Fax : +44(0)1603450045
	Email : virginie.viprey@bsrc.ac.uk
	Class : shotgun.
FEATURES	Location/Qualifiers
source	1..298 /organism="Rhizobium sp. NGR234" /strain="ANU265" /db_xref="taxon:394" /clone="04g03" /clone_lib="Shot-gun genomic library of Rhizobium strain ANU265" /note="vector: M13; derivative strain of NGR234 cured of PNGR234a"
BASE COUNT	63 a 100 c 79 g 56 t
ORIGIN	
Query Match	2.8%; Score 49.2; DB 246; Length 298;
Best Local Similarity	51.5%; Pred. No. 0.029;
Matches 139; Conservative	0; Mismatches 128; Indels 3; Gaps 1
DQY	297 GATCCAGGCCATTGCAGCCTAAACTGCGCATTCGTATCTGGGAACACCGCCCACTGCC 356 DDB 19 GCCTTACGCCATCGAATATCAAGACC GGAAAGAGCTCTGGCATTTGACAGCCGCTTGCC 78
DQY	357 GAACAATCGCCAGCCATAACAGCGTTTTGGCGAGCGCACCGCCGCGCATGGCGCTGTACGGCAC 416 DDB 79 ---CGAGGCTATCTCGCCCTGTGTGACGCTGATCAACCCCGGCGCCGCTCATCGGCCA 135 *
DQY	417 CAACGTTTACTTGTTCGTGGGACAAACACCTGTGGCGCTCGAACCGCACTGGCCA 476 DDB 136 CATATATCTATTTTCGGGACCCCTGGATGCGCCGTCTCGTGGCCCTCAACCGCAAGACGGGCCA 195
DQY	477 ACTGACGTTTCGACGTGACCGCGGCCAAGGCGAAGACATGTGTTGCACTGTCGGGCC 536 DDB 196 TGTGCTCTGGAACAAGAAGATCGCAATTAACAAGGAAGCTACAGCTACACGGCGGCC 255
DQY	537 GATCGMGCAACGCGCGTGAATGCTTGCCGG 566 DDB 256 CTTGATCTGGAAGCGCTCATCATCACCGG 285
RESULT 2	
CNS0091P	925 bp DNA GSS 03-JUN-1999
LOCUS	
DEFINITION	Drosophila melanogaster genome survey sequence TEN3 end of BAC # BCAR19D16 of RPCT-98 library from Drosophila melanogaster (fruit

RESULT	1
LOCUS	A2577200
DEFINITION	A2577200 298 bp DNA GSS 08-DEC-2000 04903 Shot-gun genomic library of Rhizobium sp. NGR234 genomic clone 04903, DNA sequence.
ACCESSION	A2577200
VERSION	A2577200.1 GI:11603441
KEYWORDS	GSS.
SOURCE	Rhizobium sp. NGR234. Rhizobium sp. NGR234
ORGANISM	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group; Rhizobiaceae; Rhizobium. 1 (bases 1 to 298)
REFERENCE	Vaiprey,V., Rosenthal,A., Broughton,W.J. and Perret,X. Genetic snapshots of the Rhizobium species NGR234 genome GenomeBiology.com 1 (6), 0014.1-0014.7 (2000) Contact: Virginie Vaiprey
AUTHORS	
TITLE	
JOURNAL	
COMMENT	

JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Séquençage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segre@genoscope.cns.fr
COMMENT - Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Oosawa and
Aaron Mammot in Pieter de Jong's laboratory in the Department of

REFERENCE 1 (bases 1 to 667)
 AUTHORS Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, J., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterson, R. and Wilson, R.
 TITLE Mashu zebrafish EST Project 1998
 JOURNAL Unpublished (1998)
 COMMENT Contact: Stephen L. Johnson
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: zbratfish@wustl.edu
 Library constructed by: Susan E. Brockerhoff DNA sequencing by: Washington University Genome Sequencing Center Clone distribution: Ressourcenzentrum/PrimaDatenbank, Berlin, Germany (web address: www.rzpd.de)
 Seq primer: T3 ET from Amersham
 High quality sequence stop: 415.

FEATURES
 source
 1..667
 /organism="Danio rerio"
 /strain="wild-type"
 /db_xref="taxon:7955"
 /clone="4199212"
 /clone_1lb="zebrafish adult retina cDNA"
 /sex="mixed"
 /dev_stage="1-2 years"
 /lab_host="E.Coli XL1-Blue MRF' (XL1-Blue MRF')"
 /note="Vector: Lambda ZAP II (pbluescript SK-); Site_1: EcoRI; Site_2: SalI; This zebrafish library was constructed by Dr. Susan E. Brockerhoff (email: sbrocker@u.washington.edu) RZPD library number: 760"

BASE COUNT 168 a 180 c 189 g 130 t
 ORIGIN

Query Match 2.6%; Score 45.6; DB 175; Length 667;
 Best Local Similarity 48.5%; Pred. No. 0.28;
 Matches 126; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

OY 366 CACGCTGAACAGCTTTGGCGAGCGACCGCGGCGATGCGCTTACGGCAACACGTTTA 425
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 391 CCCGCTGGCGCTCTCGTGGGATGAGACCTGCGAGCGCGCGAAGCTACGTCGGC 450
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 426 CTTGCTTTCGTGGGACAACCACTGGTGCCTCGACACCGCAACTGGCCCAAGTACGTT 485
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 451 GTGCGAGGTTTGGACAACATCTGCTCATCTTACAACTGAAGACCGCGAGGGGAAGT 510
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 486 CGAGCTGACCGCGCGCAAGCGAAGACATGTTTCGAACTGTCGCGGCCGATCGTGC 545
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 511 GCGCTCAGCGCGCGAGCTAGCGGACACAGAGTTATCTGCTGCTGCGCTTTCCTGGA 570
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 546 AAACGGCGTATCGTTCGGGTTTCGACCTGCAATACTCGCCGTTGGCTGCTTGTCTC 605
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 571 TGATTAACCAAGATTGACACTCAGCGACACCACTGTGCTGTGTGGACATCGAGAC 630
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 606 GGGCCACACATCGGCACCG 625
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 631 TGGCCAGCAGACGACCG 650
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
 CDS0091P 925 bp DNA GSS 03-JUN-1999
 LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
 DEFINITION BACR19D16 of RPI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
 ACCESSION AL053013
 VERSION AL053013.1 GI:4934461
 KEYWORDS GSS.

SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 925)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail: seqref@genoscope.cns.fr)
 COMMENT - Web : www.genoscope.cns.fr
 Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Oseogawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
 source
 1..925
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone_1lb="RPI-98"
 /clone="BACR19D16"
 /note="end : TET3"

BASE COUNT 120 a 61 c 61 g 172 t 511 others
 ORIGIN

Query Match 2.6%; Score 45.4; DB 219; Length 925;
 Best Local Similarity 13.2%; Pred. No. 0.33;
 Matches 40; Conservative 140; Mismatches 122; Indels 0; Gaps 0;

OY 520 TCGAAGCTGCGGGGCGCGGATCGTGCAAGCGGCTGTCGCTGCACTGCA 579
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 896 BCSSCSBSSBSSSTSMSSBSSBSSSSSGSSSGTSACVKMASSCCCGGMBACM 837
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 580 TACTCGCGCTTCGCGCTTGTCTCGGCGCACGACTCGCGCACGGTGAAGCTGTGG 639
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 836 CSSSSSCGGSASARGVKVRASGAGAKRGSGGASASHSSSSBSSSSCSASMSASS 777
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 640 CGCACTACTTCAATCCCGCGCGCTGGCGGAAGAGGTGATGAGACTTGGGCAAGATTAC 699
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 776 SSSASSSSRSRSGGAGGSGASSSSSSSSSASGASVSSASSSSSVSCSSVASS 717
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 700 GAAGCCGCTTGATGACCGCGTGGCTGGCGGCGAGATACCTATGACCCGCTACCAACTT 759
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 716 MSCSSBSSSASASSSSSSSASACASCCCTTSMSCSTASAMASARSSSSSSSCSS 657
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 760 GTCCACTACGCTGACCGCTGTGGTCCGCGCTCGAAGAACCAACGCGGACCCGGGC 819
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 656 SMSASSSSSSSSSSSSSSSSSSSGACGBSMSSGGSGSVASSGMSVSSSGRSS 597
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 820 GG 821
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 596 GS 595
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 6
 AM827060 616 bp mRNA EST 17-MAY-2000
 LOCUS AM827060
 DEFINITION IK57H09.y1 Zebrafish adult retina cDNA Danio rerio 5' similar to gb:X04526_cds2 GUNNINE NUCLEOTIDE-BINDING PROTEIN G(1)/G(S)/G(T) BETA SUBUNIT (HUMAN); mRNA sequence.
 ACCESSION AM827060

VERSION	AM827060.1	GI:7920137
KEYWORDS	EST	
SOURCE	zebrafish.	
ORGANISM	Danio rerio	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes; Cyprinidae; Rasbora; Danio.	
REFERENCE	Clark, M., Johnson, S.L., Lehnach, H., Lee, R., Li, F., Marra, M., Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.	
AUTHORS		
TITLE	Mashu zebrafish EST Project 1998	
JOURNAL	Unpublished (1998)	
COMMENT	Contact: Stephen L. Johnson Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: zbrafish@wustl.edu Library constructed by: Susan E. Brockerhoff DNA sequencing by: Washington University Genome Sequencing Center Clone distribution ResonancezentrumPrimaDatenbank, Berlin, Germany (web address: www.rzp.de) Seq primer: T3 ET from Amersham High quality sequence stop: 427. Location/Qualifiers	
FEATURES	1. 616 /organism="Danio rerio" /strain="wild-type" /db_xref="taxon:7955" /clone_lib="Zebrafish adult retina cDNA" /sex="mixed" /dev_stage="1-2 years" /lab_host="E.Coli XL1-Blue MRF' (XL1-Blue MRF')" /note="Vector: Lambda ZAP II (pBluescript SK-); Site-1: EcoRI; Site-2: SalI; This Zebrafish library was constructed by Dr. Susan E. Brockerhoff (email: sbrocker@washington.edu) RZP library number: 760"	
BASE COUNT	126 a 184 c 174 g 132 t	
ORIGIN		
Query Match	2.5%; Score 43.6; DB 121; Length 616;	
Best Local Similarity	48.1%; Pred. No. 0.91;	
Matches 124; Conservative	0; Mismatches 134; Indels 0; Gaps 0	
QY	366 CACCGTGAACAGCTTTGGCGAGCCGACCCGGCGATGGCGCTGTACGGCACCACACGTTTA 425	
Db	78 CCCGCTCGCTCCCTCGTGGTGTATGACCTGTGTGACGGCGCGCTCCGGGAACATACGTGGC 137	
QY	426 CTTTGTTCGGGGAGACACCACTGGTGGCCCTCCGACACCGCAACCTGGCCAAAGTACCTT 485	
Db	138 GTGGGGAGGTTTGGACACATCTGTCTCCATCTTACAACTGTAAGACCCGCGAGGGAGACGT 197	
QY	486 CGACGTGACCGCGGCGCCAAAGGAGACATGTGTTTGAATCTGTGGGCGCGGATCGTGGC 545	
Db	198 GCGGCTGACCGCGAGAGTATGCGCGGACACACAGATTATCTGTCTCTCTGTCTTCCTGGA 257	
QY	546 AAACGGCGTGTGCTGGCGGTTTGCACCTGGCAATATCTGCGCGTTTGCTCTTCCTC 605	
Db	258 TGTATTAACCAATGTGTGACCAAGCTCAGCGACACCAACCACTGTGTGTGTGGAGCATCGAGAC 317	
QY	606 GGGCGACAGATCGGCCAC 623	
Db	318 AGGCGACGACGACGACAC 335	
RESULT	7	
LOCUS	BC308084 662 bp mRNA EST 22-FEB-2001	

DEFINITION	fm56f02.y1 zebrafish adult retina cDNA Danio rerio CDNA clone 4195547.v1 similar to SW:GBHL HUMAN P04901 GUANINE NUCLEOTIDE-BINDING PROTEIN G(I)/G(S)/G(T) BETA SUBUNIT 1 ; mRNA sequence.					
ACCESSION	BG308084.1 GI:13105611					
VERSION	EST.					
KEYWORDS	zebrafish.					
SOURCE						
ORGANISM	Danio rerio Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Actinopteri; Ostariophysi; Cypriniformes; Cyprinidae; Rasbora; Danio.					
REFERENCE	Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Shuk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R., and Wilson.R. Washu Zebrafish EST Project 1998					
JOURNAL	Unpublished (1998)					
COMMENT	Contact: Stephen L. Johnson Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel.: 314 286 1800 Fax: 314 286 1810 Email: zbrafish@wustl.wustl.edu Library constructed by: Susan E. Brockerhoff DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: Ressourcecentumprimaratenbank, Berlin, Germany (web address: www.rzpd.de) Seq primer: "T3 ET from Amersham High quality sequence stop: 412. Location/Qualifiers 1..662 /organism="Danio rerio" /strain="wild-type" /db_xref="taxon:7955" /clone="4195547" /clone_lib="Zebrafish adult retina cDNA" /sex="mixed" /dev_stage="1-2 years" /lab_host="E. Coli XL1-Blue MRF' (XL1-Blue MRF')" /note="vector: Lambda Zap II (pluescript SK-)"; Site_1: EcoRI; Site_2: SalI; This Zebrafish library was constructed by Dr. Susan E. Brockerhoff (email: sbrocke@eu.washington.edu) RZPD library number: 760"					
BASE COUNT	161 a 175 c 194 g 132 t					
ORIGIN						
Query Match	2.5%; Score 43.6; DB 175; Length 662;					
Best Local Similarity	48.1%; Pred. No. 0.92;					
Matches 124; Conservative	0; Mismatches 134; Indels 0; Gaps 0;					
OY	366	CAGCGTAACAAGTTTGAGGCACGCCGCGCATGGCCGTGAAGCGCAACCTTTA	425			
Db	353	CCGCGTCGCTCCTCGTGGGTATGACTGTCGACCGCGCGTCCGGAACTACGTGGC	412			
OY	426	CTTTGTTCTGTGGACAAACCACCTGATGGCCCTGCACACCGCACTGGCCAATGACATT	485			
Db	413	GTTGGGAGGTTTGGACAACAATCTGCTCATGTACAACCTGAAGAAGCCGGAGGAACT	472			
OY	486	CGACGTCGACCGGGCCAAAGCGGAAGACATGTTTTGCACTGTCGGGCCGATCTGGC	545			
Db	473	GCGCGTCAGCGCGACACTAGCCGGACACACAGGTTATCTGTCTGCTGTCGTTCTCTGA	532			
OY	546	AAACGGCGTAGATGTTGACCGGTTGACCTGCACCAATATCCGCCGTTGGCGCTTTGTC	605			
Db	533	TGATTAACCAATGTGTACCACTCACGGGACACCACTGTGCTCTGTGGACATCGAGAC	592			
OY	606	GAGCCACGATCGGCCC 623				

	RESULT	8		
	BEG02586			
	LOCUS	HVSMEH0100C21f	mRNA	EST
	DEFINITION	740 bp mRNA EST library HVCDNA0009 (5 to 45 DAP) Hordeum vulgare cDNA clone HVSMEH0100C21F, mRA sequence.		
	ACCESSION	BEG02586		
	VERSION	BEG02586.2	GI:13190506	
	KEYWORDS			
	SOURCE	barley. Hordeum vulgare		
	ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;		
	REFERENCE	Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Fritsch,D., Yu ,T., Anderson,H., Dale,J., Henry,D., Kernodle,S., Palmer,M., Rambo 'T.', Sask,C., Schwartzbeck,J., Simmons,J., Choi,D.W., Main,D. and Wood,T. Development of a genetically and physically anchored EST resource for barley genomics Unpublished (2000) On Aug 21, 2000 this sequence version replaced gi:9860147. Contact: Wing RA Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Tel.: 864 656 7288 Fax: 864 656 4293 Email: rwing@clemson.edu Seq primer: ATTTAACCTCATCTAAAGG High quality sequence stop: 532. Location/Qualifiers		
	FEATURES			
	source	1..740 organism="Hordeum vulgare" cultivar="Morex" db_xref="taxon:4513" clone="HVSMEH0100C21f" clone.lib="Hordeum vulgare 5-45 DAP spike EST library HVCDNA0009 (5 to 45 DAP)" tissue.type="5-45 DAP Spike" lab.host="SOLR" note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI"		
	BASE COUNT	134 A 259 C 226 G 121 T		
	ORIGIN			
	Query Match	2.5%; Score 43.6; DB 137; Length 740;		
	Best Local Similarity	46.2%; Pred. No. 0.94;		
	Matches 145; Conservative	0; Mismatches 169; Indels 0; Gaps 0		
OY	240 CCNAGTCAGGCCCCATGCATGACGGCGGTCAATGTACTTGCGCAAACC GGCGACGTGAT	299		
Db	118 CAAGAAGAGGCGCACCGACGACGAGGCGTCCGCCGAAGAAGTGGCGGCTTGTCGACACATT	177		
OY	300 CCAGGCCAATGACGCCAAAATGCGCATCTGATCTGGGAACACCGCGCCAATGTCCCGAA	359		
Db	178 CTACAACCTCTGTAACCGACATCATCAAGTAGGGGCTGGGGCCAGTCCCTTCCACTTCTCGCC	237		
OY	360 CATGGCAGCGCTGAACAGCTTTGGGAGCGGACCGCGGACCTGTGATGAGCGACACAA	419		
Db	238 CTCCTCCCCCGGCGGTCCCACCGCAGCGCACCGCCGCTGCACAGAGAGAGCGCTCGCGCA	297		
OY	420 CGTTACTTTGTTGCTGGACAAACCACTGTGTGCGCCCTTCGACACCGCAAATGGCCAAGT	479		
Db	298 CCTGCTGCGGACAGCGCCGGGACACCGCGTCTCGAAGTGTGGATGGGCGCTCATCGTNGCC	357		
OY	480 GAGCTTGAGCTGACCGCGGGCCAAAGCAAGCAATGTTTTCAATCGTGTGGGCCCGAT	539		
Db	358 CATGGCGGCAATGCGCCGCTACAGGCGCCCACTTGTTGTGTGGATCAACATCAACAGGATA	417		

[illegible]

DB	151	MMCMCMSTCCMCGCCGCC	134
RESULT	10		
LOCUS	BE211954		
DEFINITION	BE211954	308 bp mRNA	EST
ACCESSION	BE211954	894024H07.Y1 C. reinhardtii CC-1690, normalized, lambda zap II	29-JUN-2000
VERSION	BE211954	Chlamydomonas reinhardtii cDNA, mRNA sequence.	
KEYWORDS	BE211954.1	GI:8828112	
SOURCE	EST.		
ORGANISM	Chlamydomonas reinhardtii.		
REFERENCE	Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;		
AUTHORS	Chlamydomonadaceae; Chlamydomonas.		
TITLE	1 (bases 1 to 308)		
JOURNAL	Grossman, A., Davies, J., Federspiel, N., Harris, E., Lefebvre, P.,		
COMMENT	McMormott, J. P., Silflow, C., Stern, D. and Surzycki, R.		
	Analyses of the Chlamydomonas reinhardtii genome: A model,		
	unicellular system for analyzing gene function and regulation in		
	vascular plants; project phase 2		
	Unpublished (2000)		
	Contact: Elizabeth H. Harris		
	DCMB Box 91000		
	Duke University		
	Durham, NC 27708-1000, USA		
	Tel: 919 613 8164		
	Fax: 919 613 8177		
	Email: chlamy@duke.edu.		
FEATURES			
SOURCE	location/Qualifiers		
	1..308		
	/organism="Chlamydomonas reinhardtii"		
	/strain="CC-1690 wild type mt+ 21gr"		
	/db_xref="taxon:3055"		
	/clone_lib="C. reinhardtii CC-1690, normalized, lambda zap		
	II"		
	/note="vector: pluescript II SK-; Site.1: EcoRI; Site.2:		
	XhoI; This library, constructed by John Davies and Jeffrey		
	McMormott, combines cDNAs from CC-1690 cells grown to		
	mid-log phase in TAP (acetate-containing) medium in the		
	light, TAP medium in the dark, HS (minimal) medium in		
	ambient levels of CO2 and HS medium bubbled with 5% CO2.		
	PolyA mRNA was purified from each sample, pooled and cDNA		
	synthesized. The cDNA was directionally cloned into lambda		
	zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites.		
	pluescript II SK- plasmids were excised from the lambda		
	zap clones by superinfection with Exsist (Stratagene)		
	phage. The library was normalized using method 4 described		
	in Bonaldo et al (1996) Genome Research 6: 791-806."		

	RESULT	11	
	LOCUS	BF259495	
	DEFINITION	BF259495 759 bp mRNA EST 23-FEB-2001 HVSMEF0019H19f Hordeum vulgare seedling root EST library HVCNDA0007 (etiolated and unstressed) Hordeum vulgare cDNA clone	
	ACCESSION	HVSMEF0019H19f, mRNA sequence.	
	VERSION	BF259495	
	KEYWORDS	BF259495.2 GI:13120022	
SOURCE	EST.		
ORGANISM	Barley.		
	Hordeum vulgare		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae ; Triticeae; Hordeum. 1 (bases 1 to 759)		
REFERENCE	Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D., Yu, Y., Anderson,H., Dale,J., Henry,D., Kernodle,S., Palmer,M., Rambo T., Saski,C., Schwartzbeck,J., Simmons,J., Choi,D.W., Malin,D. and Wood,T.		
AUTHORS	Development of a genetically and physically anchored EST resource for barley genomics		
	Unpublished (2000) On Nov 16, 2000 this sequence version replaced gi:1188608.		
JOURNAL	Contact: Wing RA		
COMMENT	Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29654, USA Tel.: 864 656 7288 Fax: 864 656 4293 Email: twing@clemson.edu Seq primer: AATTACCCTCACTAAAGG High quality sequence stop: 757. Location/Qualifiers 1..759		
FEATURES			
SOURCE			

Query Match	2.5%	Score 43	DB 146	Length 759
Best Local Similarity	42.5%	Pred. No. 1.3		
Matches 223	Conservative	0	Mismatches 302	Indels 0
			Gaps	0
QY	CGCGGTGGGAAACCCACGCGGCGGCCGCGGCGCTGTACGACGACGACACCGCTT	847		
Db	53 CGCGGACGACGGGGCGTTCCTCAACATGTCTGTCAACCTCATGTGGGCCCAAGAACCA	112		
QY	848 TCGCGGTGCGTCCCTGACGCGGCGAGATTGTCTGGCGTCAACACACCCCTCCCGGACA	907		
Db	113 TGGAGATCGGCGTCAACGGGCTACTCCGTGTGGCACCGGCGTGGCATCCCGAGG	172		
QY	908 ACTGGGACCAAGGATGACGTTGAGATGATGTACCAATGTGATGTCCAAACCTTCA	967		
Db	173 ACGGACCACTCTTGGCCATGACATCAACCGCGAAGACTACGAGCTGGGGCTCCCTGGA	232		
QY	968 CCGAGATGGAAGGTCTGCAGTGCATCAACCCGAAGCGCCGAACCTGGGCGACGTCGGGTGC	1027		
Db	233 TCGAGGAAGCGCGGCGTGGGGCACAAGATGCACCTTCGGAGAGGCCGCGCTCCCGCTCC	292		
QY	1028 TGACCGGCGTCCGTGCAAAACCGGACCATGTGGCACTTGACGCCGAAACCGGGAAT	1087		

QY 825 GCTGACGCGACGACACCCGTTTCGGGCTGCTCTGACACGCGGAGATTGCTGGCG 884
 ||||| ||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 Db 283 GCTGTTGGCCGACGACTACTGTTTCAAGCTGATTCACACGAGATGGGATATCTGGG 342
 QY 885 TGACACAGACCTGCCCCGCGACACTGGAGACGAGAAATGACAGTTCAGATGATGTCAC 944
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 343 CCACATATCCCGGACACATGCTGCTCTGAGATGAGATGAGATGAGAGGAGAAAGACAC 402
 QY 945 CAATGTGATGTCACACCTGACAGGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 1004
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 403 GTTTAGAGATGACCTGACAGGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 462
 QY 1005 CGCAATGCGAGAGCTGCTGCTGACCGGCTGCTGCAAAACCGGACCATGATGCA 1064
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 463 CCGTGACAGAGAGCGTTTGTCTGCCAGTATCTGTTCAGGCGACGACATTTGCCA 522
 QY 1065 GTTGCAC 1071
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 523 GGTGCGC 529

RESULT 14

BE364747 553 bp mRNA EST 20-JUL-2000
 LOCUS P11_15_A12.B1_A002 Pathogen induced 1 (P11) sorghum bicolor cDNA,
 DEFINITION mRNA sequence.
 ACCESSION BE364747
 VERSION BE364747.1 GI:9306202
 KEYWORDS EST.

SOURCE

ORGANISM Sorghum.
 Sorghum bicolor
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Sorghum.
 1 (bases 1 to 553)
 Cordonnier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt
 ,L.H.

TITLE An EST database from Sorghum: pathogen-induced plants
 JOURNAL Unpublished (2000)
 COMMENT Contact: Cordonnier-Pratt MM
 Department of Botany
 The University of Georgia
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 542 1805
 Email: mmpratt@uga.edu

REFERENCE Sequences have been trimmed to exclude PolyA, vector and regions
 below phred quality 16. The threshold for highest quality sequence
 is 20.
 Seq primer: JEN REV
 High quality sequence stop: 553
 POLYA-No.

FEATURES

source

1.553 Location/Qualifiers
 /organism="Sorghum bicolor"
 /db_xref="taxon:4558"
 /clone_lib="Pathogen induced 1 (P11)"
 /note="Organ: Anthracnose-infected leaves from
 two-week-old sorghum plants 48 hr after inoculation;
 Vector: pBluescript II from Lambda Zap II; Site.1: XhoI;
 Site.2: EcoRI; Two-week-old sorghum plants (BRX 623
 cultivar) were infected with pathogen (isolate PM421 of
 Colletotrichum graminiicola, which is a sorghum isolate).
 RNA was prepared from infected leaves harvested from 45
 seedlings 48 hours after inoculation. Note: young
 seedlings (2 weeks old) exhibit juvenile resistant
 reaction, which is an incompatible interaction. As they
 grow older (4 weeks or older), plants resume susceptibility
 to anthracnose disease. The library was made from poly-A
 RNA in the cloning vector lambda zap II. Clones to be
 sequenced were prepared by mass excision. WARNING: While
 most or all ESTs are expected to derive from the host
 plant, no effort was made to eliminate ESTs deriving from

BASE COUNT 98 a 199 c 182 g 73 t 1 others
 ORIGIN

Query Match 2.4%; Score 42; DB 166; Length 553;
 Best Local Similarity 46.6%; Pred. No. 2.3; Mismatches 155; Indels 0; Gaps 0;
 Matches 135; Conservative 0;

QY 1443 TGTCTGCGGCGACACTATTCGCGGCTTGTGACCGGCGGCGGCTTGTTCACAGGTG 1502
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 56 TGTCTGCGGCGACACTATTCGCGGCTTGTGACCGGCGGCGGCTTGTTCACAGGTG 115
 QY 1503 TAGGATGCTTACTTCGCGGCTTGTGACCGGCGGCGGCTTGTTCACAGGTG 1562
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 116 ACGCACCTCCACTTCTACATGACGACGAGTACACCGGCGGAGAACCCGACGCGCGCT 175
 QY 1563 CCTTGCACCGCTGCGGCGGCGGCGGCGGCTTGTGACCGGCGGCGGCTTGTTCACAGGTG 1622
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 176 CATCATGCGAGGAGAGGCGGCGGCTTGTGACCGGCGGCGGCTTGTTCACAGGTG 235
 QY 1623 CGCATGCGAGGAGGAGGCGGCGGCTTGTGACCGGCGGCGGCTTGTTCACAGGTG 1682
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 236 CGACGACGAGGCGGCGGCGGCTTGTGACCGGCGGCGGCTTGTTCACAGGTG 295
 QY 1683 GCGAGTGCATGACCGGCGGCTTGTGACCGGCGGCGGCTTGTTCACAGGTG 1732
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 296 GCCCGAGCGGCGGCGGCGGCTTGTGACCGGCGGCGGCTTGTTCACAGGTG 345

RESULT 15

BG557222 469 bp mRNA EST 10-APR-2001
 LOCUS EML_44.B11.g1_A002 Embryo 1 (EML) Sorghum bicolor cDNA, mRNA
 DEFINITION sequence.
 ACCESSION BG557222
 VERSION BG557222.1 GI:13586220
 KEYWORDS EST.

SOURCE

ORGANISM Sorghum.
 Sorghum bicolor
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Sorghum.
 1 (bases 1 to 469)
 Reid,S.P., Cordonnier-Pratt,M.-M., Gingle,A. and Pratt,L.H.

REFERENCE An EST database from Sorghum: developing embryos
 JOURNAL Unpublished (2000)
 COMMENT Contact: Cordonnier-Pratt MM
 Department of Botany
 The University of Georgia
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 542 1805
 Email: mmpratt@uga.edu

REFERENCE Sequences have been trimmed to exclude PolyA, vector and regions
 below phred quality 16. The threshold for highest quality sequence
 is 20.
 Seq primer: PolyMix
 High quality sequence start: 2
 High quality sequence stop: 440
 POLYA-No.

FEATURES

source

1.469 Location/Qualifiers
 /organism="Sorghum bicolor"
 /db_xref="taxon:4558"
 /clone_lib="Embryo 1 (EML)"
 /note="Organ: Embryos germinated for 24 hr; Vector:
 pBluescript II from Lambda Zap II; Site.1: XhoI; Site.2:
 EcoRI; The library was made from poly-A RNA in the cloning
 vector lambda zap II. Clones to be sequenced were
 prepared by mass excision."
 BASE COUNT 87 a 122 c 156 g 104 t
 ORIGIN

